FORM PTO-1390 (REV. 10-2000) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY'S DOCKET NUMBER
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		B0192/7031
		U S APPLICATION NO (If known, see 37 CFR 1 5) 09/890229
INTERNATIONAL APPLICATION NO. PCT/GB00/00263	INTERNATIONAL FILING DATE 28 January 2000 (28.01.00)	PRIORITY DATE CLAIMED 28 January 1999 (28.01.99)

TITLE OF INVENTION
MANIPULATING ISOPRENOID EXPRESSION

APPLICANT(S) FOR DO/EO/US
BRAMLEY, Peter Michael; HARKER, Mark

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☒ This is an express request to promptly begin national examination procedures (35 U.S.C. 371(f)).
4. ☒ The US has been elected by the expiration of 19 months from the earliest claimed priority date (PCT Article31).
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2)).
   a. ☐ is transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☒ has been transmitted by the International Bureau.
   c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
   a. ☐ is transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☐ has been transmitted by the International Bureau.
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3)).
   a. ☐ are attached hereto (required only if not transmitted by the International Bureau).
   b. ☐ have been communicated by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☒ have not been made and will not be made.
8. ☐ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☐ An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(C)(5)).

Items 11. To 16. Below concern document(s) or information included:

11. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. ☒ A FIRST preliminary amendment.
14. ☐ A SECOND or SUBSEQUENT preliminary amendment.
15. ☐ A substitute specification.
16. ☐ A change of power of attorney and/or address letter.
17. ☒ A computer-readable form of the sequence listing in accordance with PCT Rule 13ter.2 and 35 U.S.C.1.821-1.825.
18. ☐ A second copy of the published international application under 35 U.S.C. 154(d)(4).
19. ☐ A second copy of the English language translation of the international application under 35 U.S.C. 154(d)(4).
20. ☒ Other items or information:

Copy of PCT/RO/101 form
Copy of PCT Published Application with Search Report
PCT/IB/308 form

Copy of International Preliminary Examination Report
Statement Pursuant to 37 CFR 1.821(f) w/ diskette and paper copy of sequence listing
Express Mail Label No. EL819461955US
Date Mailed: July 27, 2001

JC17 Rec'd PCT/PTO 27 JUL 2001

U.S. APPLICATION NO (if known, see 37 CFR 1.5) 09/890229	INTERNATIONAL APPLICATION PCT/GB00/00263	ATTORNEY'S DOCKET NUMBER B0192/7031

21. ☒ The following fees are submitted:

		CALCULATIONS PTO USE ONLY	
BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5)):			
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO	$1000.00		
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO	$860.00		
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but but international search fee paid to USPTO (37 CFR 1.445(a)(2)).paid to USPTO	$710.00		
International preliminary examination fee paid to USPTO (37 CFR 1.482) But all claims did not satisfy provisions of PCT Article 33(1)-(4)	$690.00		
International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4)	$100.00		
ENTER APPROPRIATE BASIC FEE AMOUNT = 860.00			
Surcharge of $130.00 for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(e)).		$	

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total Claims	34-20 =	14	X $18.00	$252.00	
Independent Claims	6- 3 =	3	X $80.00	$240.00	
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+$270.00	$	
TOTAL OF ABOVE CALCULATIONS			=	$1352.00	
Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by ½.				$	
SUBTOTAL			=	$1352.00	
Processing fee of $130.00 for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(f)).				$	
TOTAL NATIONAL FEE			=	$1352.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate coversheet (37 CFR 3.28, 3.31). $40.00 per property +				$	
TOTAL FEES ENCLOSED			=	$1352.00	
				Amount to be: refunded	$
				charged	$

a. ☒ A check in the amount of $ 1352.00 To cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. _____ In the amount of $_____ To cover the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 23/2825. A duplicate of this sheet is enclosed.

d. ☐ Fees are to be charged to a credit card. WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO

WOLF, GREENFIELD & SACKS, P.C.
600 Atlantic Avenue
Boston, Massachusetts 02210
Tel: (617) 720-3500

CUSTOMER NUMBER 23628

John R. Van Amsterdam
SIGNATURE

John R. Van Amsterdam
NAME

40,212
REGISTRATION NO

495341.1

JC17 Rec'd PCT/PTO 27 JUL 2001

U.S. APPLICATION NO (If known, see 37 CFR 1.5) 09/890229	INTERNATIONAL APPLICATION PCT/GB00/00263	ATTORNEY'S DOCKET NUMBER B0192/7031

21. ☒ The following fees are submitted:

BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5)):

		CALCULATIONS PTO USE ONLY
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO	$1000.00	
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO	$860.00	
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but but international search fee paid to USPTO (37 CFR 1.445(a)(2)).paid to USPTO	$710.00	
International preliminary examination fee paid to USPTO (37 CFR 1.482) But all claims did not satisfy provisions of PCT Article 33(1)-(4)	$690.00	
International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4)	$100.00	
ENTER APPROPRIATE BASIC FEE AMOUNT = 860.00		

Surcharge of $130.00 for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(e)).				$	
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total Claims	34-20 =	14	X $18.00	$252.00	
Independent Claims	6- 3 =	3	X $80.00	$240.00	
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+$270.00	$	
TOTAL OF ABOVE CALCULATIONS			=	$1352.00	
Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by ½.				$	
SUBTOTAL			=	$1352.00	
Processing fee of $130.00 for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(f)).				$	
TOTAL NATIONAL FEE			=	$1352.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate coversheet (37 CFR 3.28, 3.31). $40.00 per property +				$	
TOTAL FEES ENCLOSED			=	$1352.00	
				Amount to be: refunded	$
				charged	$

a. ☒ A check in the amount of $ 1352.00 To cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. _____ In the amount of $_____ To cover the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 23/2825. A duplicate of this sheet is enclosed.

d. ☐ Fees are to be charged to a credit card. WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO.

WOLF, GREENFIELD & SACKS, P.C.
600 Atlantic Avenue
Boston, Massachusetts 02210
Tel: (617) 720-3500

CUSTOMER NUMBER 23628

John R. Van Amsterdam
SIGNATURE

John R. Van Amsterdam
NAME

40,212
REGISTRATION NO

495341.1

Attorney's Docket No. **B0192/7031**

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants	:	Bramley et al.
U.S. Serial No.	:	09/890,229
International Application No.	:	PCT/GB00/00263
International Filing Date	:	28 January 2000 (28.01.00)
Earliest Priority Date	:	28 January 1999 (28.01.99)
Title	:	MANIPULATING ISOPRENOID EXPRESSION
Examiner	:	Unknown
Art Unit	:	1651

CERTIFICATE OF EXPRESS MAILING UNDER 37 C.F.R. §1.8(a)

The undersigned hereby certifies that this document is being placed in the United States mail via express mail, addressed to Box PCT, U.S. Patent and Trademark Office, P.O. Box 2327, Arlington, VA 22202, on the 27th day of November, 2001.

Monica E. Zombori

Box PCT
U.S. Patent and Trademark Office
P.O. Box 2327
Arlington, VA 22202

COMMUNICATION

Dear Sir:

The Notification of Missing Requirements which was mailed on September 6, 2001, states that claims 12, 13, and 26 are multiple dependent claims, and as such the multiple dependent claim fee of $270.00 is required for these claims. However, this seems to be an error. Upon review of the application, Applicants found that claims 12, 13, and 26 were not multiple dependent claims. Therefore, Applicants believe there is no additional claim fee required in this application.

If any fee is determined to be required, please charge the balance to the account of the undersigned, deposit account number 23/2825.

Respectfully submitted,

John R. Van Amsterdam, Reg. No. 40,212
WOLF, GREENFIELD & SACKS, P.C.
600 Atlantic Avenue
Boston, MA 02210-2211
Tel. (617) 720-3500

Attorney's Docket No.: B0192/7031
Date: November 27, 2001
x12/06/01

577963.1

CRF Errors Corrected by the STIC Systems Branch

Serial Number: 09/870229

CRF Processing Date: 10/05/2001
Edited by: mm
Verified by: _____ (STIC sta

☐ Changed a file from non-ASCII to ASCII

☐ Changed the margins in cases where the sequence text was "wrapped" down to the next line.

☐ Edited a format error in the Current Application Data section, specifically:

_____

☐ Edited the Current Application Data section with the actual current number. The number inputted by the applicant was ☐ the prior application data; or ☐ other _____

ENTERED

☐ Added the mandatory heading and subheadings for "Current Application Data".

☐ Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.

☐ Changed the spelling of a mandatory field (the headings or subheadings), specifically:

_____

☐ Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were:

_____

☐ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

_____

☐ Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.

☐ Inserted colons after headings/subheadings. Headings edited included:

_____

☐ Deleted extra, invalid, headings used by an applicant, specifically:

_____

☑ Deleted: ☑ non-ASCII "garbage" at the beginning/end of files; ☐ secretary initials/filename at end of file; ☐ page numbers throughout text; ☐ other invalid text, such as _____

☐ Inserted mandatory headings, specifically: _____

☐ Corrected an obvious error in the response, specifically:

_____

☐ Edited identifiers where upper case is used but lower case is required, or vice versa.

☐ Corrected an error in the Number of Sequences field, specifically:

_____

☐ A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.

☐ Deleted *ending* stop codon in amino acid sequences and adjusted the "(A)Length:" field accordingly (error due to a PatentIn bug). Sequences corrected: _____

☐ Other:

_____

_____

_____

_____

*Examiner: The above corrections must be communicated to the applicant in the first Office Action. DO NOT send a copy of this form.

3/1/95

#3

Attorney's Docket No. **B0192/7031**

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants	:	Bramley et al.
U.S. Serial No.	:	09/890,229
International Application No.	:	PCT/GB00/00263
International Filing Date	:	28 January 2000 (28.01.00)
Earliest Priority Date	:	28 January 1999 (28.01.99)
Title	:	MANIPULATING ISOPRENOID EXPRESSION
Examiner	:	Unknown
Art Unit	:	1651

CERTIFICATE OF EXPRESS MAILING UNDER 37 C.F.R. §1.8(a)

The undersigned hereby certifies that this document is being placed in the United States mail via express mail, addressed to Box PCT, U.S. Patent and Trademark Office, P.O. Box 2327, Arlington, VA 22202, on the 27th day of November, 2001.

Monica E. Zombori

Box PCT
U.S. Patent and Trademark Office
P.O. Box 2327
Arlington, VA 22202

STATEMENT PURSUANT TO 37 CFR 1.821(f) AND 37 CFR 1.825 (a) and (b)

This statement is made pursuant to 37 CFR 1.821 (f), and 37 CFR 1.825 (a) and (b). Applicants submit herewith a substitute copy of the written sequence listing and a computer readable diskette to comply with the sequence requirements.

Applicants' attorney states that the information recorded in computer readable form is identical to the written sequence listing and that neither the computer readable form nor the written sequence listing contain new matter.

Respectfully submitted,

John R. Van Amsterdam, Reg. No. 40,212
WOLF, GREENFIELD & SACKS, P.C.
600 Atlantic Avenue
Boston, MA 02210-2211
Tel. (617) 720-3500

Attorney's Docket No.: B0192/7031
Date: November 27, 2001
x12/06/01

577949.1

Attorney's Docket No: B0192/7031

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants	:	Peter BRAMLEY et al.
International Application No.	:	PCT/GB00/00263
International Filing Date	:	28 January 2000 (28.01.00)
Earliest Priority Date	:	28 January 1999 (28.01.99)
Title	:	MANIPULATING ISOPRENOID EXPRESSION

Commissioner for Patents
Box PCT
Washington, DC 20231

PRELIMINARY AMENDMENT

Sir:

Please amend the United States national phase application of the above-identified PCT international application as follows.

In the Specification

Please amend the specification as follows:

Please add the following section as the first section of the specification following the title.

Related Applications

This application is a national stage filing under 35 U.S.C. § 371 of PCT International application PCT/GB00/00263, and filed 28 January 2000, which was published under PCT Article 21(2) in English.

Foreign priority benefits are claimed under 35 U.S.C. § 119(a)-(d) of Great Britain application number GB 9901902.8, filed 28 January 1999.

Please substitute paragraphs of the specification as follows.

Substitute the following paragraph for the paragraph that begins on line 33, page 13 of the specification as filed.

Figure 3: is an illustration of an amino acid sequence alignment of *DXP* synthases used in the present invention, Synechocystis sp. 6803 (S.s) (GenBank D90903) (SEQ ID NO: 1), B. subtilis (B.s) (GenBank D84432) (SEQ ID NO: 2) and *E. coli* (E.c) (GenBank AF035440) (SEQ ID NO: 3). The consensus line (consen) shows residues conserved in all three sequences (upper case letters) or residues which are identical in two sequences and replaced by an equivalent amino acid in the third sequence (+). The conserved histidine domain putatively involved in proton transfer is over lined and numbered 1. The second over lined domain (2) denotes the consensus thiamin pyrophosphate (TPP)-binding motif.

Substitute the following paragraph for the paragraph that begins on line 35, page 14 of the specification as filed.

Figure 6: is a diagrammatic illustration of vector pVB6_TSEC_LML (SEQ ID NO: 6).

Substitute the following paragraph for the paragraph that begins on line 3, page 15 of the specification as filed.

Figure 7: is a diagrammatic representation of plasmid pVB6_35S_TSEC_LML (SEQ ID NO: 5).

Substitute the following paragraph for the paragraph that begins on line 6, page 15 of the specification as filed.

Figure 8: is an illustration of the amino acid sequence of *E. coli* DXPS (SEQ ID NO: 3).

Substitute the following paragraph for the paragraph that begins on line 9, page 15 of the specification as filed.

Figure 9: is an illustration of the transit peptide used in tomato plants (SEQ ID NO: 4).

Substitute the following paragraph for the paragraph that begins on line 15, page 16 of the specification as filed.

Based on the nucleotide sequence of ORF s111945 from the genome database for Synechocystis sp. PCC 6803 [23], primers were designed to clone the putative dxps gene by polymerase chain reaction (PCR). The forward primer 5'-GTCCCAATCCACCATGCACATCAG-3' (SEQ ID NO: 7) overlaps the beginning of the coding sequence. The reverse primer 5'-CCCTCGACAAATGCAAAATGTATC-3' (SEQ ID NO: 8) lies outside the stop codon of the gene. A PCR (25 cycles) using Pfu DNA polymerase (Stratagene) produced a DNA fragment of the expected size (~1.9 kb). Subsequent sequencing of the fragment confirmed the product to be the ORF s111945. The B. subtilis dxps gene was also cloned by PCR using primers designed to amplify the gene encoding the product YqiE, identified in the Bacillus subtilis genome database [24]. The forward primer 5'-GATCCGCTATGGATCTTTTATC-3' (SEQ ID NO: 9) contains a modified base substitution at the predicted start codon (underlined) for improved expression in *E. coli*. The reverse primer 5'-ATCTAATCGTTCTTTCTTTGAC-3' (SEQ ID NO: 10) lies outside the stop codon of the dxps gene. After PCR (25 cycles) a DNA product of the expected size (~1.9 kb) was obtained, and when sequenced proved to be identical to the gene encoding the product YqiE. The PCR products from both reactions were treated with Taq DNA polymerase (GibcoBRL) at 72°C for 10 min to synthesise blunt ended fragments. The fragments were then cloned into the EcoRV site of the pBluescript vector (Stratagene) using T4 DNA ligase (Fermentas) (Fig. 2).

Substitute the following paragraph for the paragraph that begins on line 32, page 20 of the specification as filed.

The amino acid sequence of the DXPS proteins of Synechocystis sp. 6803 (SEQ ID NO: 1) and B. subtilis (SEQ ID NO: 2) exhibited significant similarity to each other over their entire length (47% identities) and to *E. coli* DXPS (SEQ ID NO: 3) (B. subtilis (44% identities) and Synechocystis sp. 6803 (46% identities)) (Fig. 3). All three polypeptides share two conserved domains; one thought to be involved in thiamin binding [30] and a histidine residue postulated to participate in proton transfer [31], both of which are detailed in Fig. 3. The existence of thiamin-

binding domain in each of the polypeptides explains the cofactor requirement of thiamin for DXPS activity [12]. The high degree of polypeptide sequence identity, particularly the distribution of conserved domains, in all three indicates that they all encode DXPS or a closely related gene product.

Substitute the following paragraph for the paragraph that begins on line 34, page 27 of the specification as filed.

Forward: 5'-GCG CCG CTA TTT ACT CGA-3' (SEQ ID NO: 11)

Substitute the following paragraph for the paragraph that begins on line 1, page 28 of the specification as filed.

Reverse: 5'-TTT CTC TGG CGT GCC GCC-3' (SEQ ID NO: 12)

A marked up copy of the specification amendments is attached as Appendix A to facilitate the Examiner's review.

Please substitute the enclosed Sequence Listing for the presently filed Sequence Listing.

In the Claims

Please amend the claims as follows:

5.(amended) A method according to claim 3 wherein said vector comprises one or more nucleic acid sequences encoding a polypeptide(s) capable of producing a desired isoprenoid.

6.(amended) A method according to claim 3 wherein said plant or plant cell is transformed with a further vector comprising one or more nucleic acid sequences encoding a polypeptide(s) capable of producing a desired isoprenoid.

7.(amended) A method according to claim 3 wherein said vector comprising said nucleic acid sequence(s) encoding said DXPS and/or said polypeptide(s) capable of producing said isoprenoid further comprises a nucleic acid sequence of either a tissue specific promoter and/or encoding a plastid transit peptide.

8.(amended) A method according to claim 5 wherein said desired isoprenoid is one conferring a nutritional benefit.

16.(amended) A method according to claim 14, wherein said cell is any of a bacterial, yeast or algal cell.

18.(amended) A method according to claim 14 wherein said organism is a plant.

19.(amended) A method according to claim 14 wherein said cell is any of a bacterial, yeast or algal cell.

20.(amended) A method according to claim 14 wherein said bacterial cell is *E. coli*.

27.(amended) A transgenic cell, tissue or organism according to claim 24, wherein said organism is a plant.

29.(amended) Progeny of the organism according to claim 24 having increased isoprenoid activity.

30.(amended) A transformed plant comprising a transgene capable of expressing DXPS from *E. coli* having the sequence according to SEQ ID NO: 3 and which plant comprises a higher level of isoprenoid than an untransformed plant.

31.(amended) A transformed plant according to claim 30 comprising any of constructs pVB6_TSEC_LML (SEQ ID NO: 6) or pVB6_35S_TSEC-LML (SEQ ID NO: 5).

32.(amended) A transformed plant according to claim 30 wherein said plant is a tomato plant.

A marked up copy of the amended claims is attached as Appendix B to facilitate the Examiner's review.

Remarks

Applicants have amended the specification to provide priority application information and information regarding the publication in English under PCT Article 21(2) of the PCT application of which the above-identified application is a U.S. national stage application. The specification also has been amended to insert sequence identifiers. A substitute Sequence Listing is filed herewith.

Applicants have amended the claims to eliminate certain dependencies and to add sequence identifiers. No new matter has been added.

Respectfully submitted,

John R. Van Amsterdam
Reg. No. 40,212
WOLF, GREENFIELD & SACKS, P.C.
600 Atlantic Avenue
Boston, Massachusetts 02210-2211
Tel: (617) 720-3500

Attorney's Docket No. B0192/7031
Dated: 27 July 2001

Appendix A

Added Section

Related Applications

This application is a national stage filing under 35 U.S.C. § 371 of PCT International application PCT/GB00/00263, and filed 28 January 2000, which was published under PCT Article 21(2) in English.

Foreign priority benefits are claimed under 35 U.S.C. § 119(a)-(d) of Great Britain application number GB 9901902.8, filed 28 January 1999.

Amended Paragraphs of the Specification

Substitute the following paragraph for the paragraph that begins on line 33, page 13 of the specification as filed.

Figure 3: is an illustration of an amino acid sequence alignment of *DXP* synthases used in the present invention, Synechocystis sp. 6803 (S.s) (GenBank D90903) (SEQ ID NO: 1), B. subtilis (B.s) (GenBank D84432) (SEQ ID NO: 2) and *E. coli* (E.c) (GenBank AF035440) (SEQ ID NO: 3). The consensus line (consen) shows residues conserved in all three sequences (upper case letters) or residues which are identical in two sequences and replaced by an equivalent amino acid in the third sequence (+). The conserved histidine domain putatively involved in proton transfer is over lined and numbered 1. The second over lined domain (2) denotes the consensus thiamin pyrophosphate (TPP)-binding motif.

Substitute the following paragraph for the paragraph that begins on line 35, page 14 of the specification as filed.

Figure 6: is a diagrammatic illustration of vector pVB6_TSEC_LML (SEQ ID NO: 6).

Substitute the following paragraph for the paragraph that begins on line 3, page 15 of the specification as filed.

Figure 7: is a diagrammatic representation of plasmid pVB6_35S_TSEC_LML (SEQ ID NO: 5).

Substitute the following paragraph for the paragraph that begins on line 6, page 15 of the specification as filed.

Figure 8: is an illustration of the amino acid sequence of *E. coli* DXPS (SEQ ID NO: 3).

Substitute the following paragraph for the paragraph that begins on line 9, page 15 of the specification as filed.

Figure 9: is an illustration of the transit peptide used in tomato plants (SEQ ID NO: 4).

Substitute the following paragraph for the paragraph that begins on line 15, page 16 of the specification as filed.

Based on the nucleotide sequence of ORF s111945 from the genome database for Synechocystis sp. PCC 6803 [23], primers were designed to clone the putative dxps gene by polymerase chain reaction (PCR). The forward primer 5'-GTCCCAATCCACCATGCACATCAG-3' (SEQ ID NO: 7) overlaps the beginning of the coding sequence. The reverse primer 5'-CCCTCGACAAATGCAAAATGTATC-3' (SEQ ID NO: 8) lies outside the stop codon of the gene. A PCR (25 cycles) using Pfu DNA polymerase (Stratagene) produced a DNA fragment of the expected size (~1.9 kb). Subsequent sequencing of the fragment confirmed the product to be the ORF s111945. The B. subtilis dxps gene was also cloned by PCR using primers designed to amplify the gene encoding the product YqiE, identified in the Bacillus subtilis genome database [24]. The forward primer 5'-GATCCGCTATGGATCTTTTATC-3' (SEQ ID NO: 9) contains a modified base substitution at the predicted start codon (underlined) for improved expression in *E. coli*. The reverse primer 5'-ATCTAATCGTTCTTTCTTTGAC-3' (SEQ ID NO: 10) lies outside

the stop codon of the dxps gene. After PCR (25 cycles) a DNA product of the expected size (~1.9 kb) was obtained, and when sequenced proved to be identical to the gene encoding the product YqiE. The PCR products from both reactions were treated with Taq DNA polymerase (GibcoBRL) at 72°C for 10 min to synthesise blunt ended fragments. The fragments were then cloned into the EcoRV site of the pBluescript vector (Stratagene) using T4 DNA ligase (Fermentas) (Fig. 2).

Substitute the following paragraph for the paragraph that begins on line 32, page 20 of the specification as filed.

The amino acid sequence of the DXPS proteins of Synechocystis sp. 6803 (SEQ ID NO: 1) and B. subtilis (SEQ ID NO: 2) exhibited significant similarity to each other over their entire length (47% identities) and to *E. coli* DXPS (SEQ ID NO: 3) (B. subtilis (44% identities) and Synechocystis sp. 6803 (46% identities)) (Fig. 3). All three polypeptides share two conserved domains; one thought to be involved in thiamin binding [30] and a histidine residue postulated to participate in proton transfer [31], both of which are detailed in Fig. 3. The existence of thiamin-binding domain in each of the polypeptides explains the cofactor requirement of thiamin for DXPS activity [12]. The high degree of polypeptide sequence identity, particularly the distribution of conserved domains, in all three indicates that they all encode DXPS or a closely related gene product.

Substitute the following paragraph for the paragraph that begins on line 34, page 27 of the specification as filed.

Forward: 5'-GCG CCG CTA TTT ACT CGA-3' (SEQ ID NO: 11)

Substitute the following paragraph for the paragraph that begins on line 1, page 28 of the specification as filed.

Reverse: 5'-TTT CTC TGG CGT GCC GCC-3' (SEQ ID NO: 12)

Appendix B

5. A method according to claim 3 [or 4] wherein said vector comprises one or more nucleic acid sequences encoding a polypeptide(s) capable of producing a desired isoprenoid.

6. A method according to claim 3 [or 4] wherein said plant or plant cell is transformed with a further vector comprising one or more nucleic acid sequences encoding a polypeptide(s) capable of producing a desired isoprenoid.

7. A method according to [any of] claim[s] 3 [to 6] wherein said vector comprising said nucleic acid sequence(s) encoding said DXPS and/or said polypeptide(s) capable of producing said isoprenoid further comprises a nucleic acid sequence of either a tissue specific promoter and/or encoding a plastid transit peptide.

8. A method according to [any of] claim[s] 5 [to 7] wherein said desired isoprenoid is one conferring a nutritional benefit.

16. A method according to claim 14 [or 15], wherein said cell is any of a bacterial, yeast or algal cell.

18. A method according to claim 14 [or 15] wherein said organism is a plant.

19. A method according to [any of] claim[s] 14 [to 18] wherein said cell is any of a bacterial, yeast or algal cell.

20. A method according to [any of] claim[s] 14 [to 19] wherein said bacterial cell is *E. coli*.

27. A transgenic cell, tissue or organism according to [any of] claim[s] 24 [to 26], wherein said organism is a plant.

29. Progeny of the organism according to [any of] claim[s] 24 [to 28] having increased isoprenoid activity.

30. A transformed plant comprising a transgene capable of expressing DXPS from *E. coli* having the sequence according to [Figure 8] SEQ ID NO: 3 and which plant comprises a higher level of isoprenoid than an untransformed plant.

31. A transformed plant according to claim 30 comprising any of constructs pVB6_TSEC_LML (SEQ ID NO: 6) or pVB6_35S_TSEC-LML (SEQ ID NO: 5) [illustrated in Figures 6 and 7].

32. A transformed plant according to claim 30 [or 31] wherein said plant is a tomato plant.

SEQUENCE LISTING

<110> Bramley, Peter Michael
Harker, Mark

<120> Manipulating Isoprenoid Expression

<130> B0192/7031

<140> 09/890,229
<141> 2000-01-28

<150> GB 9901902.8
<151> 1999-01-28

<160> 12

<170> PatentIn version 3.0

<210> 1
<211> 640
<212> PRT
<213> Synechocystis sp.

<400> 1

Met His Ile Ser Glu Leu Thr His Pro Asn Glu Leu Lys Gly Leu Ser
1 5 10 15

Ile Arg Glu Leu Glu Glu Val Ser Arg Gln Ile Arg Glu Lys His Leu
20 25 30

Gln Thr Val Ala Thr Ser Gly Gly His Leu Gly Pro Gly Leu Gly Val
35 40 45

Val Glu Leu Thr Val Ala Leu Tyr Ser Thr Leu Asp Leu Asp Lys Asp
50 55 60

Arg Val Ile Trp Asp Val Gly His Gln Ala Tyr Pro His Lys Met Leu
65 70 75 80

Thr Gly Arg Tyr His Asp Phe His Thr Leu Arg Gln Lys Asp Gly Val
85 90 95

Ala Gly Tyr Leu Lys Arg Ser Glu Ser Arg Phe Asp His Phe Gly Ala
100 105 110

Gly His Ala Ser Thr Ser Ile Ser Ala Gly Leu Gly Met Ala Leu Ala
115 120 125

Arg Asp Ala Lys Gly Glu Asp Phe Lys Val Val Ser Ile Ile Gly Asp
130 135 140

Gly Ala Leu Thr Gly Gly Met Ala Leu Glu Ala Ile Asn His Ala Gly
145 150 155 160

His Leu Pro His Thr Arg Leu Met Val Ile Leu Asn Asp Asn Glu Met
165 170 175

Ser Ile Ser Pro Asn Val Gly Ala Ile Ser Arg Tyr Leu Asn Lys Val

```
                180                 185                 190

Arg Leu Ser Ser Pro Met Gln Phe Leu Thr Asp Asn Leu Glu Glu Gln
        195                 200                 205

Ile Lys His Leu Pro Phe Val Gly Asp Ser Leu Thr Pro Glu Met Glu
    210                 215                 220

Arg Val Lys Glu Gly Met Lys Arg Leu Val Val Pro Lys Val Gly Ala
225                 230                 235                 240

Val Ile Glu Glu Leu Gly Phe Lys Tyr Phe Gly Pro Ile Asp Gly His
                245                 250                 255

Ser Leu Gln Glu Leu Ile Asp Thr Phe Lys Gln Ala Glu Lys Val Pro
            260                 265                 270

Gly Pro Val Phe Val His Val Ser Thr Thr Lys Gly Lys Gly Tyr Asp
        275                 280                 285

Leu Ala Glu Lys Asp Gln Val Gly Tyr His Ala Gln Ser Pro Phe Asn
    290                 295                 300

Leu Ser Thr Gly Lys Ala Tyr Pro Ser Ser Lys Pro Lys Pro Pro Ser
305                 310                 315                 320

Tyr Ser Lys Val Phe Ala His Thr Leu Thr Thr Leu Ala Lys Glu Asn
                325                 330                 335

Pro Asn Ile Val Gly Ile Thr Ala Ala Met Ala Thr Gly Thr Gly Leu
            340                 345                 350

Asp Lys Leu Gln Ala Lys Leu Pro Lys Gln Tyr Val Asp Val Gly Ile
        355                 360                 365

Ala Glu Gln His Ala Val Thr Leu Ala Ala Gly Met Ala Cys Glu Gly
    370                 375                 380

Ile Arg Pro Val Val Ala Ile Tyr Ser Thr Phe Leu Gln Arg Gly Tyr
385                 390                 395                 400

Asp Gln Ile Ile His Asp Val Cys Ile Gln Lys Leu Pro Val Phe Phe
                405                 410                 415

Cys Leu Asp Arg Ala Gly Ile Val Gly Ala Asp Gly Pro Thr His Gln
            420                 425                 430

Gly Met Tyr Asp Ile Ala Tyr Leu Arg Cys Ile Pro Asn Leu Val Leu
        435                 440                 445

Met Ala Pro Lys Asp Glu Ala Glu Leu Gln Gln Met Leu Val Thr Gly
    450                 455                 460

Val Asn Tyr Thr Gly Gly Ala Ile Ala Met Arg Tyr Pro Arg Gly Asn
465                 470                 475                 480

Gly Ile Gly Val Pro Leu Met Glu Glu Gly Trp Glu Pro Leu Glu Ile
                485                 490                 495

Gly Lys Ala Glu Ile Leu Arg Ser Gly Asp Asp Val Leu Leu Leu Gly
```

```
            500                 505                 510

Tyr Gly Ser Met Val Tyr Pro Ala Leu Gln Thr Ala Glu Leu Leu His
        515                 520                 525

Glu His Gly Ile Glu Ala Thr Val Val Asn Ala Arg Phe Val Lys Pro
    530                 535                 540

Leu Asp Thr Glu Leu Ile Leu Pro Leu Ala Glu Arg Ile Gly Lys Val
545                 550                 555                 560

Val Thr Met Glu Glu Gly Cys Leu Met Gly Gly Phe Gly Ser Ala Val
                565                 570                 575

Ala Glu Ala Leu Met Asp Asn Asn Val Leu Val Pro Leu Lys Arg Leu
            580                 585                 590

Gly Val Pro Asp Ile Leu Val Asp His Ala Thr Pro Glu Gln Ser Thr
        595                 600                 605

Val Asp Leu Gly Leu Thr Pro Ala Gln Met Ala Gln Asn Ile Met Ala
    610                 615                 620

Ser Leu Phe Lys Thr Glu Thr Glu Ser Val Val Ala Pro Gly Val Ser
625                 630                 635                 640
```

<210> 2
<211> 633
<212> PRT
<213> Bacillus subtilis

<400> 2

```
Met Asp Leu Leu Ser Ile Gln Asp Pro Ser Phe Leu Lys Asn Met Ser
1               5                   10                  15

Ile Asp Glu Leu Glu Lys Leu Ser Asp Glu Ile Arg Gln Phe Leu Ile
            20                  25                  30

Thr Ser Leu Ser Ala Ser Gly Gly His Ile Gly Pro Asn Leu Gly Val
        35                  40                  45

Val Glu Leu Thr Val Ala Leu His Lys Glu Phe Asn Ser Pro Lys Asp
    50                  55                  60

Lys Phe Leu Trp Asp Val Gly His Gln Ser Tyr Val His Lys Leu Leu
65                  70                  75                  80

Thr Gly Arg Gly Lys Glu Phe Ala Thr Leu Arg Gln Tyr Lys Gly Leu
                85                  90                  95

Cys Gly Phe Pro Lys Arg Ser Glu Ser Glu His Asp Val Trp Glu Thr
            100                 105                 110

Gly His Ser Ser Thr Ser Leu Ser Gly Ala Met Gly Met Ala Ala Ala
        115                 120                 125

Arg Asp Ile Lys Gly Thr Asp Glu Tyr Ile Ile Pro Ile Ile Gly Asp
    130                 135                 140
```

```
Gly Ala Leu Thr Gly Gly Met Ala Leu Glu Ala Leu Asn His Ile Gly
145                 150                 155                 160

Asp Glu Lys Lys Asp Met Ile Val Ile Leu Asn Asp Asn Glu Met Ser
                165                 170                 175

Ile Ala Pro Asn Val Gly Ala Ile His Ser Met Leu Gly Arg Leu Arg
            180                 185                 190

Thr Ala Gly Lys Tyr Gln Trp Val Lys Asp Glu Leu Glu Tyr Leu Phe
        195                 200                 205

Lys Lys Ile Pro Ala Val Gly Gly Lys Leu Ala Ala Thr Ala Glu Arg
    210                 215                 220

Val Lys Asp Ser Leu Lys Tyr Met Leu Val Ser Gly Met Phe Phe Glu
225                 230                 235                 240

Glu Leu Gly Phe Thr Tyr Leu Gly Pro Val Asp Gly His Ser Tyr His
                245                 250                 255

Glu Leu Ile Glu Asn Leu Gln Tyr Ala Lys Lys Thr Lys Gly Pro Val
            260                 265                 270

Leu Leu His Val Ile Thr Lys Lys Gly Lys Gly Tyr Lys Pro Ala Glu
        275                 280                 285

Thr Asp Thr Ile Gly Thr Trp His Gly Thr Gly Pro Tyr Lys Ile Asn
    290                 295                 300

Thr Gly Asp Phe Val Lys Pro Lys Ala Ala Ala Pro Ser Trp Ser Gly
305                 310                 315                 320

Leu Val Ser Gly Thr Val Gln Arg Met Ala Arg Glu Asp Gly Arg Ile
                325                 330                 335

Val Ala Ile Thr Pro Ala Met Pro Val Gly Ser Lys Leu Glu Gly Phe
            340                 345                 350

Ala Lys Glu Phe Pro Asp Arg Met Phe Asp Val Gly Ile Ala Glu Gln
        355                 360                 365

His Ala Ala Thr Met Ala Ala Ala Met Ala Met Gln Gly Met Lys Pro
    370                 375                 380

Phe Leu Ala Ile Tyr Ser Thr Phe Leu Gln Arg Ala Tyr Asp Gln Val
385                 390                 395                 400

Val His Asp Ile Cys Arg Gln Asn Ala Asn Val Phe Ile Gly Ile Asp
                405                 410                 415

Arg Ala Gly Leu Val Gly Ala Asp Gly Glu Thr His Gln Gly Val Phe
            420                 425                 430

Asp Ile Ala Phe Met Arg His Ile Pro Asn Met Val Leu Met Met Pro
        435                 440                 445

Lys Asp Glu Asn Glu Gly Gln His Met Val His Thr Ala Leu Ser Tyr
    450                 455                 460
```

Attorney's Docket No: B0192/7031

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants	:	Peter BRAMLEY et al.
International Application No.	:	PCT/GB00/00263
International Filing Date	:	28 January 2000 (28.01.00)
Earliest Priority Date	:	28 January 1999 (28.01.99)
Title	:	MANIPULATING ISOPRENOID EXPRESSION

Commissioner for Patents
Box PCT
Washington, D.C. 20231

STATEMENT UNDER 37 C.F.R. §1.821(f)

Sir:

This statement is made pursuant to 37 CFR 1.821(f). Applicants enclose herewith an original written copy of the Sequence Listing and a computer readable diskette. Applicants' attorney states that the information recorded in the computer readable form is identical to the written Sequence Listing and that the Sequence Listing contains no new matter.

Respectfully submitted,

John R. Van Amsterdam, Reg No. 40,212
Wolf, Greenfield & Sacks, P.C.
600 Atlantic Avenue
Boston, MA 02210
Telephone (617) 720-3500

Docket No. B0192/7031
Dated: July 27, 2001
xNDD

MANIPULATING ISOPRENOID EXPRESSION

The present invention is concerned with manipulating or altering isoprenoid expression in a cell or organism which biosynthesises isopentenyl diphosphate (IPP), which is the universal precursor of all isoprenoids in nature, via a mevalonate independent pathway.

Isoprenoids constitute the largest class of natural products occurring in nature, with over 29,000 individual compounds identified to date [1]. Chemically, they are extremely diverse in their structure and complexity. The fundamental biological functions performed by isoprenoids ensure they are essential for the normal growth and developmental processes in all living organisms. These include functioning as eukaryotic membrane stabilisers (sterols), plant hormones (gibberellins and abscisic acid), providing pigments for photosynthesis (carotenoids and phytol side chain of chlorophyll), and as carriers for electron transport (menaquinone, plastoquinone and ubiquinone).

All isoprenoids are synthesised via a common metabolic precursor, isopentenyl diphosphate (IPP; C_5). Until recently, the biosynthesis of IPP was generally assumed to proceed exclusively from acetyl-CoA via the classical mevalonate pathway (Fig. 1) [2]. The enzyme 3-hydroxy-3-methylglutaryl coenzyme A reductase (*HMGR*) catalyses the conversion of hydroxymethylglutaryl-CoA to mevalonate, a key reaction of the mevalonate-dependant IPP biosynthetic pathway. Recent studies have demonstrated that mevalonate is not the biosynthetic precursor of IPP in all living organisms

[3,4]. The existence of an alternative, mevalonate-independent pathway for IPP formation was identified initially in several species of eubacteria [4,5] and a green alga [6]. The pathway was subsequently shown to be operational in the plastids of higher plants [7-10]. The first reaction in the non-mevalonate pathway is the transketolase-type condensation reaction of pyruvate and D-glyceraldehyde-3-phosphate to yield 1-deoxy-D-xylulose-5-phosphate (DXP) (Fig. 1). This reaction is catalysed by the enzyme 1-deoxy-D-xylulose-5-phosphate synthase. The second reaction in the pathway is the conversion of DXP to 2-C-methyl-D-erythritol-4-phosphate (MEP). The reactions which convert MEP to IPP have yet to be characterised.

The cloning and characterisation of the DXP synthase (*dxps*) gene has been described for a number of organisms including Esherichia coli [11,12] and higher plants [13-15]. The *CLA1* gene product from *Arabidopsis thaliana* associated with chloroplast development [16], for example, has been shown to exhibit *DXPS* activity [11]. Recently, a gene responsible for the reduction of DXP to 2-C-methyl-D-erythritol-4-phosphate, the proposed next step in the non-mevalonate pathway has been cloned from *E. coli* [17].

The present inventors have surprisingly found that the first reaction in the mevalonate-independent IPP biosynthetic pathway is highly influential in controlling the levels of isoprenoids which can be formed in a cell or organism within which the mevalonate independent IPP biosynthetic pathway is present. The enzyme DXPS or functional equivalents

thereof, has been identified by the present inventors as a rate-limiting step in isoprenoid biosynthesis and that DXPS activity plays an important role in channelling the carbon resources of the cell into the isoprenoid biosynthetic pathway.

Therefore, according to a first aspect of the present invention there is provided a method of manipulating isoprenoid expression in a cell possessing a mevalonate independent isopentenyl diphosphate synthesising pathway, which method comprises altering the activity of the enzyme 1-deoxy-D-xylulose-5-phosphate synthase (DXPS), or a functional equivalent thereof. Thus, advantageously, the rate-limiting effect conferred by DXPS activity on the IPP biosynthetic pathway can be utilised to manipulate the resultant levels of isoprenoids in a cell by altering the activity or expression of DXPS.

Preferably, the levels of isoprenoids in a cell can be enhanced by increasing the activity or expression of the DXPS. Likewise reduced levels of isoprenoids can be achieved by reducing or inhibiting activity or expression of DXPS in a cell or organism. Increasing the DXPS activity may be achieved by, for example, transforming the cell which may itself be part of a cell line or an organism, with an expression vector comprising a nucleic acid molecule encoding DXPS, which may advantageously be operably linked to a reporter molecule, such as used in the GUS assay which is known in the art. Preferably, the vector comprises any of the vectors designated pBSDXPS or pSYDXPS, illustrated in Figure 2.

An alternative method for altering expression may comprise utilising a technique known as Enforced Evolution, or DNA Shuffling see Patten *et al.* Current Opinion in Biotechnology, 1997, Vol. 8, No. 6, pp 724-733, Crameri *et al.*, Nature 1998, Vol. 391, No. 6664, pp 288-291 and Harayama S, Trends in Biotechnology, 1998, Vol. 16, No. 2, pp 76-82. According to this method improvements in enzyme activity can be achieved by reassembling DNA segments into a full length gene by homologous or site specific combination. Before the assembly, the segments are often subjected to random mutagenesis by error prone PCR, random nucleotide insertion, or other such methods. The genes can be expressed in suitable microbial hosts leading to the production of functional polypeptides, such as DXPS.

The nucleic acid encoding the DXPS may be endogenous to the cell or organism into which it will be transformed or, alternatively, may be exogenous. In one embodiment of the invention, the method may also comprise transforming the cell or organism with a vector comprising one or more nucleic acid sequences suitable for producing a desired isoprenoid. This aspect of the invention is particularly advantageous because it allows isoprenoids to be produced in a cell or organism independent of the source of the isoprenoid which may be derived from cells or organisms which do not possess the mevalonate independent IPP biosynthesising pathway. Similarly, enhanced levels of an isoprenoid can be produced in cells or organisms having the mevalonate independent IPP biosynthetic pathway.

Therefore, in the example where the cell is *E. coli* it is possible to engineer production of an isoprenoid which is exogenous to the *E. coli* bacterium, which isoprenoid may be, for example, any of the carotenoids of plants, such as, lycopene or even an isoprenoid of human origin.

Carotenoids are yellow-orange-red lipid based pigments found in nature. They have been found to be useful in a variety of applications, for example, as supplements, and particularly vitamin supplements, as vegetable oil based food products and food ingredients, as feed additives in animal feeds and as colorants. Phytoene has been found to be useful in treating skin disorders whilst lycopene and α and β carotene consumption have been implicated as having preventative effects against certain kinds of cancers. Therefore, it is a highly advantageous aspect of the invention that increased production of such compounds can be achieved and which compounds can confer considerable health care benefits. Once the desired carotenoid or other isoprenoid has been produced in *E. coli*, or other suitable organism as defined above, it can be isolated using standard bicengineering techniques.

Increases in concentrations of any desired isoprenoid may be achieved, in a cell or alternatively an organism which possesses the IPP biosynthetic mevalonate independent pathway. For example, crops can be engineered using the method of the invention to produce increased levels of an isoprenoid which confers nutritional benefits to humans following consumption of the plant, such as, for example, vitamin E and lycopene.

Therefore, there is also provided by a further aspect of the invention a cell or organism having a mevalonate independent IPP biosynthetic pathway and which has been transformed or transfected with an expression vector comprising a nucleic acid molecule encoding DXPS or a functional equivalent or bioprecursor thereof. As described above, the vector may also include one or more further nucleic acid sequences suitable for producing a desired isoprenoid, or alternatively the one or more nucleic acid sequences may be included in a separate vector, operably linked to suitable expression control sequences. In a particularly preferred embodiment the cell or organism comprises a plant.

An expression vector according to the invention includes a vector having a nucleic acid sequence operably linked to regulatory sequences, such as promoter regions, that are capable of effecting expression of said DNA fragments. The term "operably linked" refers to a juxta position wherein the components described are in a relationship permitting them to function in their intended manner. Such vectors may be transformed into a suitable host cell or organism to produce a desired protein, such as DXPS or an isoprenoid according to the method of the invention. Thus, in a further aspect, the invention provides a process for producing a desired isoprenoid which comprises cultivating a host cell, transformed or transfected with an expression vector as described above under conditions to provide for expression by the vector of DXPS or a functional equivalent thereof or suitable polypeptides for producing a desired isoprenoid and optionally recovering the expressed polypeptides.

The vectors may be, for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression of said nucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable markers, such as, for example, ampicillin resistance.

Regulatory elements required for expression include promoter sequences to bind RNA polymerase and transcription initiation sequences for ribosome binding. For example, a bacterial expression vector may include a promoter such as the lac promoter and for transcription initiation in the Shine-Dalgarno sequence and the start codon AUG. Similarly, a eukaryotic expression vector may include a heterologous or homologous promoter for RNA polymerase II, a downstream polyadenylation signal, the start codon AUG, and a termination codon for detachment of the ribosome. Such vectors may be obtained commercially or assembled from the sequences described by methods well known in the art.

By combining the nucleic acid sequences encoding said DXPS and optionally the one or more sequences suitable for producing an isoprenoid with tissue specific promoters, enhanced isoprenoid levels in specified tissues of plants can be achieved. For example, by utilising a seed specific promoter or other transcriptional initiation region, elevated levels of carotenoids in seeds may be achieved. The seed can then be harvested and which provides a reservoir for the isoprenoid or carotenoid of interest.

Generally, the nucleic acid molecule encoding said DXPS which is included in the vector used in

accordance with the method of the invention, will be transformed into a plant cell so that the DXPS molecule is directed to the plastids of the plant. Accordingly, where the vector is not inserted directly into the plastid of the plant, the vector will further comprise a nucleic acid sequence operably linked to said DXPS or said one or more isoprenoid producing nucleic acid sequences and which further sequence will encode a transit peptide to direct expression of the DXPS or isoprenoid into the plastid. Native or heterologous transit peptides may be utilised in this embodiment of the invention.

As aforesaid, the mevalonate independent IPP biosynthetic pathway is not present in any higher animals, particularly humans. Therefore, the inhibition of the reaction catalysed by DXPS provides a unique target site to selectively inhibit or alleviate bacterial associated infections by altering the expression level of or inhibiting function or activity of DXPS.

One method of inhibiting or preventing expression of DXPS utilises antisense technology. Antisense technology can be used to control gene expression through helix formation of antisense DNA or RNA, both of which methods are based on polynucleotide binding to DNA or RNA. For example, the 5'-coding region of a native DNA sequence coding for DXPS according to the invention may be used to design an antisense RNA nucleotide of from 10 to 50 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple-helix - see Lee *et al*, Nucl. Acids. Res., 6:3073 (1979); Cooney *et al.*, Science, 241:456 (1988);

and Derman *et al.*, Science 251:1360 (1991), in which case expression of the antisense RNA oligonucleotide allows hybridisation to the mRNA *in vivo* and blocks translation of an mRNA molecule into DXPS.

Alternatively, compounds can be screened for their ability to inhibit the catalytic activity or expression of DXPS in the mevalonate - independent IPP biosynthetic pathway. According to a further aspect of the invention, therefore, there is also provided a method of identifying a compound which modulates isoprenoid production or expression which method comprises contacting said compound to be tested with a molecule from the mevalonate independent IPP biosynthetic pathway and which molecule undergoes a reaction in the presence of an appropriate reactant catalysed by DXPS, in the presence of DXPS and monitoring the level of product produced when compared to the same reaction in the absence of the compound to be tested. Preferably, the molecules which are reacted are pyruvate and glyceraldehyde-3-phosphate, and which undergo a condensation reaction in the presence of DXPS, to yield 1-deoxy-D-xylulose-5-phosphate (DXP) as illustrated in Figure 1.

Any compounds identified as preventing expression or activity of the DXPS enzyme according to the invention may advantageously be particularly useful as selective toxicity agents to destroy, for example, bacterial or plant cells which possess the mevalonate independent IPP biosynthetic pathway. These compounds therefore can be particularly useful as medicaments or herbicides, or alternatively in the preparation of a medicament for treating bacterial associated diseases.

A further aspect of the invention therefore also comprises a pharmaceutical composition comprising a compound identified as an inhibitor of expression or activity of DXPS or a functional equivalent or bioprecursor thereof, together with a pharmaceutically acceptable carrier, diluent or excipient thereof. Also provided by the invention is a herbicidal composition comprising said compound identified as an inhibitor of expression or activity of DXPS function.

An even further aspect of the invention comprises a transgenic cell, tissue or organism having a mevalonate independent IPP biosynthetic pathway, which comprises a transgene capable of expressing at least one additional DXPS molecule according to the invention. The transgenic cell, tissue or organism may also comprise a transgene having one or more nucleic acid sequences capable of producing a desired isoprenoid. Preferably, the transgenic cell comprises a plant and even more preferably tomato plants.

The term "transgene capable of expression" as used herein means a suitable nucleic acid sequence(s) which leads to expression of DXPS or proteins having the same function and/or activity and/or encoding proteins capable of producing a desired isoprenoid. The transgene, may include, for example, isolated genomic nucleic acid or synthetic nucleic acid, including DNA integrated into the genome. Preferably, the transgene comprises the nucleic acid sequence(s) encoding the DXPS enzyme or said isoprenoid as described herein, or a functional fragment of said nucleic acid. A functional fragment of said nucleic acid should be taken to mean a fragment of the gene comprising said nucleic acid(s) coding for the DXPS enzyme or said

isoprenoid or a functional equivalent, derivative or a non-functional derivative such as a dominant negative mutant, or bioprecursor thereof. For example, it would be readily apparent to persons skilled in the art that nucleotide substitutions or deletions may be made using routine techniques, which do not affect the protein sequence and subsequent functioning of the DXPS enzyme and/or isoprenoid producing proteins encoded by said nucleic acid(s).

The DXPS enzyme expressed or the isoprenoid produced by said transgenic cell, tissue or organism or a functional equivalent or bioprecursor of said protein also forms part of the present invention.

The recombinant DNA molecules or vectors of the invention can be introduced into a plant cell in a number of recognised ways in the art and it will be appreciated that the choice of method used might depend on the type of plant, i.e. monocot or dicot, targeted for transformation. Suitable methods of transforming plant cells include microinjection (Crossway *et al.* (1986) BioTechnieques 4:320-334), electroporation (Riggs *et al.* (1986) Proc. Natl. Acad. Sci. USA 83:5602-5606, Agrobacterium mediated transformation (Hinchee et al. (1988) Biotechnology 6:915-921) and ballistic particle acceleration (see, for example, Sanford *et al.*, U.S. Patent 4,945,050; and McCabe *et al.* (1988) Biotechnology 6:923-926).

Alternatively, in the case of an organism, such as a plant, a plastid can be transformed directly. Stable transformation of chloroplasts has been reported in higher plants, see, for example, SVAB *et al.* (1990) Proc. Natl. Acad. Sci. USA 87:8526-8530; SVAB & Maliga

(1993) Proc. Natl. Acad. Sci. USA 90:913-917; Staub & Maliga (1993) Embo J. 12:601-606. The method relies on particle gun delivery of DNA containing a selectable marker and targeting of the DNA to the plastid genome through homologous recombination. In such methods, plastid gene expression can be accomplished by use of a plastid gene promoter or by trans-activation of a silent plastid-borne transgene positioned for expression from a selective promoter sequence such as that recognised by T7 RNA polymerase. The silent plastid gene is activated by expression of the specific RNA polymerase from a nuclear expression construct and targeting of the polymerase to the plastid by use of a transit peptide. Tissue-specific expression may be obtained in such a method by use of a nuclear-encoded and plastid-directed specific RNA polymerase expressed from a suitable plant tissue specific promoter. Such a system has been reported in McBride *et al.* (1994) Proc. Natl. Acad. Sci., USA 91:7301-7305.

The cells which have been transformed may be grown into plants in accordance with conventional methods known in the art. See, for example, McCormick *et al.*, Plant Cell Reports (1986), 5:81-84. These plants may then be grown, and either pollinated with the same transformed strainer or different strains, and the resulting hybrid having the desired phenotypic characteristic identified. Two or more generations may be grown to ensure that the subject phenotypic characteristic is stably maintained and inherited and then seeds harvested to ensure the desired phenotype or other property has been achieved.

A host cell of any plant variety may be employed.

Plant species which provide seeds of interest are particularly useful. For the most part, plants will be chosen where the seed is produced in high amounts, a seed-specific product of interest is involved, or the seed or a seed part is edible. Seeds of interest include the oil seeds, such as oilseed Brassica seeds, cotton seeds, soybean, safflower, sunflower, coconut, palm, and the like; grain seeds, e.g. wheat, barley, oats amaranth, flax, rye, triticale, rice, corn, etc.; other edible seeds or seeds with edible parts including pumpkin, squash, sesame, poppy, grape, mung beans, peanut peas, beans, radish, alfalfa, cocoa, coffee, tree nuts such as walnuts, almonds, pecans, chick-peas etc.

The invention may be more clearly understood from the following exemplary embodiment described with reference to the accompanying drawings wherein:

Figure 1: is an illustration of the mevalonate-dependant (A) and independent (B) pathways for IPP biosynthesis. Proposed reactions for the biosynthesis of 1-deoxy-D-xylulose-5-phosphate from pyruvate and glyceraldehyde-3-phosphate, catalysed by *DXPS* is as shown inside the box.

Figure 2: is an illustration of structure of plasmids pBSDXPS and pSYDXPS.

Figure 3: is an illustration of an amino acid sequence alignment of *DXP* synthases used in the present invention,

Synechocystis sp. 6803 (S.s) (GenBank D90903), B. subtilis (B.s) (GenBank D84432) and *E. coli* (E.c) (GenBank AF035440). The consensus line (consen) shows residues conserved in all three sequences (upper case letters) or residues which are identical in two sequences and replaced by an equivalent amino acid in the third sequence (+). The conserved histidine domain putatively involved in proton transfer is over lined and numbered 1. The second over lined domain (2) denotes the consensus thiamin pyrophosphate (TPP)-binding motif.

Figure 4: is a graphic representation of lycopene accumulation in recombinant *E. coli* cultures expressing vector only (□), B. subtilis DXPS (•) and Synechocystis sp. 6803 DXPS (△). (Data are means ± S.E.M. from three independent determinations.)

Figure 5: is an illustration of lycopene (open columns) and UQ-8 (shaded columns) content of *E. coli* control cultures (vector only) or expressing exogenous B. subtilis dxps (B. subtilis), Synechocystis sp. 6803 dxps (sp. 6803) or A. thaliana hmgr1 (HMGR1) genes. (Data are means (S.E.M. from three independent determinations.)

Figure 6: is a diagrammatic illustration of

vector pVB6_TSEC_LML.

Figure 7: is a diagrammatic representation of plasmid pVB6_35S_TSEC-LML.

Figure 8: is an illustration of the amino acid sequence of *E.coli* DXPS.

Figure 9: is an illustration of the transit peptide used in tomato plants.

EXAMPLE 1

Materials and methods

Bacterial strains, plasmids, and culture conditions.

E. coli strain XL1-Blue (Stratagene) was used for gene cloning and expression of plasmids. *E. coli* was grown in Luria Broth media [18] at 37°C on a rotary shaker at 250 rpm (unless otherwise stated). Ampicillin (100 μg/ml), chloramphenicol (50 μg/ml) and 1.0 mM isopropyl-b--D-thiogalactoside (IPTG) (all purchased from Sigma) were added as required. Plasmid pBluescript (Stratagene) was used as a vector for both cloning and expression studies. Synechocystis sp. PCC 6803 was obtained from the Institute Pasteur (Paris) and grown in BG11 medium [19] supplemented with 0.5% glucose at 30°C and 2,000 lux. Bacillus subtilis strain PY79 DNA was a kind gift from P. Wakeley (Royal Holloway, University of London). The construction of plasmid pACCRT-EIB, which expresses the E. uredovora crtE, crtB and crtI genes necessary for lycopene biosynthesis in *E. coli* cells into which it has been

introduced, has been described previously [20]. The plasmid used for the expression of HMGR1 cloned into pBluescript (pHMGR1) has also been described elsewhere [21].

Recombinant DNA techniques

All recombinant DNA techniques were performed by standard methods [22] or according to suppliers instructions. Genomic DNA was extracted from all organisms using the Qiagen Genomic-tip 20/G kit.

Cloning of dxps genes

Based on the nucleotide sequence of ORF sll1945 from the genome database for Synechocystis sp. PCC 6803 [23], primers were designed to clone the putative dxps gene by polymerase chain reaction (PCR). The forward primer 5'-GTCCCAATCCACCATGCACATCAG-3' overlaps the beginning of the coding sequence. The reverse primer 5'-CCCTCGACAAATGCAAAATGTATC-3' lies outside the stop codon of the gene. A PCR (25 cycles) using Pfu DNA polymerase (Stratagene) produced a DNA fragment of the expected size (~1.9 kb). Subsequent sequencing of the fragment confirmed the product to be the ORF sll1945. The B. subtilis dxps gene was also cloned by PCR using primers designed to amplify the gene encoding the product YqiE, identified in the Bacillus subtilis genome database [24]. The forward primer 5'-GATCCGCTATGGATCTT TTATC-3' contains a modified base substitution at the predicted start codon (underlined) for improved expression in *E. coli*. The reverse primer 5'-ATCTAATCGTTCTTTCTTTGAC-3' lies outside the stop codon of the dxps gene. After PCR (25 cycles) a DNA product of the expected size (~1.9 kb) was

obtained, and when sequenced proved to be identical to the gene encoding the product YqiE. The PCR products from both reactions were treated with Taq DNA polymerase (GibcoBRL) at 72°C for 10 min to synthesise blunt ended fragments. The fragments were then cloned into the EcoRV site of the pBluescript vector (Stratagene) using T4 DNA ligase (Fermentas) (Fig. 2).

In vitro DXP synthase assay

E. coli XL1-blue cells, transformed with the appropriate plasmid, were grown at 37°C in Luria Broth medium with appropriate antibiotics to an $OD_{620\ nm}$ of 0.6, and induced by the addition of 1.0 mM IPTG at 28°C for two hours. Bacteria were harvested by centrifugation (6,000g for 10 min) and washed in buffer A (100 mM Tris-HCl (pH 7.5), 1 mM dithioreitol, 0.3 M sucrose). Cells were resuspended to their original volume in buffer B (100 mM Tris (pH 8.0), 1 mM dithioreitol, 0.1 mM phenylmethanesulphonyl fluoride, 1μ/ml pepstatin, 1μg/ml leupeptin, 1 mg/ml lysozyme). The cells were then incubated at 30°C for 15 min with gentle agitation, and then disrupted by brief sonication at 4°C. The supernatant was recovered and the protein concentration determined using the Bradford assay [25].

An aliquot of the supernatant (100 μl) was transferred to an Eppendorf tube along with 100 μl of assay buffer containing 100 mM Tris (pH 8.0), 3 mM ATP, 3 mM Mn^{2+}, 3 mM Mg^{2+}, 1 mM KF, 1 mM thiamine diphosphate, (0.1%) Tween 60, 0.6 Mm mDL-glyceraldehyde-3-phosphate, 30 μM [2-^{14}C]pyruvate (0.5 μCi). The mixture was incubated for 2 hours at

30°C with gentle agitation. The reaction was stopped by heating the mixture at 80°C for 3 min. After centrifugation at 13,000 g for 5 min, the supernatant was transferred to a clean tube and evaporated to dryness. The residue was resuspended in methanol (50 μl) and loaded onto a TLC plate (silica gel 60). Chromatograms were developed in *n*-propyl alcohol/ethyl acetate/H_2O (6:1:3 v/v/v).

Enzyme assays were performed with extracts of induced cells expressing either Synechocystis sp. PCC 6803 or B. subtilis DXPS, as opposed to control assays in which cells contained only the pBluescript vector without insert. TLC analysis of assays expressing one of the dxps clones exhibited a major band (R_f 0.14) assumed to be DXP which was not observed in the controls. Quantification of ^{14}C-labelled DXP was achieved by isolation of the reaction product on TLC. The DXP band was scraped off the plate, eluted from the silica using methanol and quantified by liquid-scintillation counting. Enzymatic dephosphorylation of the assay products resulted in the formation of 1-deoxy-D-xylulose (DX), when analysed on TLC (R_f 0.50). When non-radioactive pyruvate was used in the assay, the DXP (Rf 0.12 stained purple) and DX (R_f 0.47 stained blue) were identified by staining with p-anisaldehyde/sulphuric acid (3:1). The DXP co-chromatographed with authentic, chemically synthesised DXP which stained purple also. The reaction substrates pyruvate (R_f 0.36 stained yellow), DL-glyceraldehyde-3-phosphate (R_f 0.15 stained orange) and D-glyceraldehyde (R_f 0.74 stained orange) were also observable using this TLC system. In reactions where the assay products were dephosphorylated no DXP was observed on TLC only DX.

Quantification of lycopene and ubiquinone QB-8 in *E. coli*

Bacterial growth was determined from the $OD_{620\ nm}$. Dry cell weight was calculated from known volumes of culture harvested by centrifugation at 13,000 g for 5 min, washed once with water and recentrifuged. The cells were lyophilised overnight and the weight of the dried cell pellet determined. The lycopene content of the cells was determined by harvesting aliquots of *E. coli* cells by centrifugation at 13,000 g for 5 min and washing once in water followed by recentrifuging. The cells were resuspended in acetone (200 μl) and incubated at 68°C for 5 min in the dark. The samples were centrifuged again 13,000 g for 10 min and the acetone supernatant containing the lycopene was placed in a clean tube. The extract was evaporated to dryness under a stream of N2 and stored at -20°C in the dark. The lycopene content of the extracts was determined by visible light absorption spectra using a Beckman DU Series 7000 diode array spectrometer. Spectra were recorded in acetone using an $A^{1\%}_{1cm}$ of 3450 [26].

UQ-8 was extracted from cells based on the methods of Yoshida et al. [27]. Cells were collected by centrifugation, washed once with water and then lyophilised overnight. The dried pellet was extracted in *n*-propanol (3 ml) and of *n*-hexane (5 ml) containing 15 μg of UQ-10 as an internal standard, by disruption of the cells using a pestle and mortar. The solvent phase and that obtained by the second extraction from the aqueous phase n-hexane (3 ml) were combined and evaporated to dryness under N_2. The residue was resuspended in ethanol and analysed by reversed phase

HPLC as described previously [28]. Peaks were identified by comparing their elution profiles with standards for UQ-7, UQ-9 and UQ-10. A standard of UQ-8 was not available, and the UQ-8 peak was identified by its elution profile relative to those of the other standards [29].

Cloning of the dxps genes

The cloning of dxps and the characterisation of the gene product, DXPS, from *E. coli* has recently been reported by two research groups [11,12]. The gene product was shown to exhibit DXP synthase activity, which is considered as the first reaction of the mevalonate-independent pathway for IPP biosynthesis (Fig. 1) [5]. Based on the *E. coli* dxps nucleotide sequence homologs of the gene were identified in the eubacterial genomes of B. subtilis and Synechocystis sp. PCC 6803. The open reading frame sll1945 in the Synechocystis sp. 6803 genome was cloned by PCR, ligated into the vector pBluescript, and designated pSYDXPS (Fig. 2). The gene extends over 1920 bp and contains an open reading frame encoding a polypeptide of 640 amino acids, with a predicted molecular mass of 69 kDa. The dxps homolog in the B. subtilis genome was identified as the ORF encoding the product YqiE. It was cloned by PCR, and introduced into pBluescript to generate plasmid pBSDXPS (Fig. 2). The gene extends over 1899 bp and encodes a polypeptide of 633 amino acids with a predicted molecular mass of 70 kDa.

The amino acid sequence of the DXPS proteins of Synechocystis sp. 6803 and B. subtilis exhibited significant similarity to each other over their entire length (47% identities) and to the *E. coli* DXPS (B.

subtilis (44 % identities) and Synechocystis sp. 6803 (46 % identities)) (Fig. 3). All three polypeptides share two conserved domains; one thought to be involved in thiamin binding [30] and a histidine residue postulated to participate in proton transfer [31], both of which are detailed in Fig. 3. The existence of a thiamin-binding domain in each of the polypeptides explains the cofactor requirement of thiamin for DXPS activity [12]. The high degree of polypeptide sequence identity, particularly the distribution of conserved domains, in all three indicates that they all encode DXPS or a closely related gene product.

Quantification of lycopene and UQ-8 in *E. coli* transformants

Cells of *E. coli* transformed with pACCRT-EIB [20] are pigmented pink due to the accumulation of lycopene. *E. coli* cells engineered to produce lycopene, were transformed with either pBSDXPS, pSYDXPS, pHMGR, or pBluescript to act as a control, to monitor the effect on lycopene biosynthesis when exogenous DXPS was expressed in the cells. The *E. coli* were grown in 50 ml cultures at 30°C with induction by IPTG for 48 hours, by which time they had reached the stationary phase of growth. Figure 4 shows the accumulation of lycopene in the cultures during the 48 hour culture period. The graph clearly demonstrates that the *E. coli* cultures expressing exogenous dxps accumulated lycopene at a much greater rate than the control culture. The final lycopene content of the recombinant dxps strains was approximately double that of the control (Fig. 5). A similar increase was also obtained in *E. coli* cells engineered to produce the

this organism.

In vitro enzyme activity

The increased levels of carotenoids and UQ-8 in *E. coli* expressing exogenous DXPS were hypothesised to be due to increased DXPS enzymatic activity in the cells. This was confirmed by preparing cell homogenates from recombinant *E. coli* strains after induction with IPTG. Reaction products were measured over a two hour period, separated by TLC and quantified by liquid-scintillation counting. The major product obtained from the reaction co-chromatographed with chemically-synthesised DXP. This confirms DXP as the major reaction product in the assay. The putative DXPS function of B. subtilis ORF encoding the product YqiE and Synechocystis sp. 6803 ORF sll1945 has been established by these results. Table 1 shows the specific activity of DXPS in the recombinant *E. coli* strains. The results show that DXPS activity was increased in *E. coli* expressing endogenous dxps genes. This increase was greatest in homogenates containing the B. subtilis DXPS, where a 2.0 fold increase was observed compared to the controls. Homogenates containing the Synechocystis sp. 6803 DXPS exhibited a 1.8 fold increase compared to control reactions. Therefore, increased DXPS activity in *E. coli* appears to be responsible for the increased levels of carotenoids and UQ-8 observed in the transgenic strains. The relative increases in carotenoid levels between *E. coli* cultures expressing plasmids pSYNDXSP and pBSDXPS closely resemble the increases observed in the in vitro studies. This suggests that there is a direct relationship between DXPS activity and the carotenoid content of the cells. This is not the case

colourless carotenoid phytoene (data not shown). Alterations in the endogenous levels of isoprenoids were determined by measuring the ubiquinone content of the cells. In *E. coli*, the major quinones encountered are ubiquinone (UQ-8) and menaquinone (MK-8) [32]. Ubiquinone is a major component of the aerobic respiratory chain. It is estimated that there are approximately 50 molecules of ubiquinone for each of the oxidation complexes in *E. coli* [33]. By measuring an end product which is produced in relatively large quantities, it was conjectured that alterations in the rates of biosynthesis could be readily detected. The UQ-8 content of the recombinant dxps strains was 1.5 times greater than the controls (Fig. 5). Lycopene and UQ-8 levels were measured in *E. coli* transformed with hmgr1 from A. thaliana, to monitor if this caused any alterations in the isoprenoid content of the cells. Expression of the A. thaliana hmgr1 cDNA had no effect of the levels of lycopene nor UQ-8 in the cells (Fig. 5).

The results show that increased expression of DXPS leads to increased lycopene and UQ-8 levels in the recombinant *E. coli* cells. This indicates that increasing the rate of DXP synthesis, the initial reaction in the mevalonate-independent pathway for IPP biosynthesis, elevates isoprenoid production. In contrast, expression of hmgr1 had no effect on isoprenoid biosynthesis, suggesting that mevalonate dependent IPP biosynthesis has little or no role in IPP synthesis in *E. coli*. Similarity searches of the *E. coli* genome data base for proteins of the mevalonate-dependent IPP biosynthesis pathway failed to identify any possible homologs in the genome suggesting that this pathway is probably absent in

for UQ-8 where increases in the levels of UQ-8 are more restricted, which could be due to a rate-limiting reaction later in the UQ-8 biosynthesis pathway [34]. The results support the hypothesis that increased DXPS activity in *E. coli* results in increased levels of carotenoids and UQ-8. These data suggest that isoprenoid levels in *E. coli* can be increased by enhancing DXPS activity.

Transformation Protocols in Tomato Plants

Triparental mating

Liquid LB medium (5ml) containing rifampicin (100μg/ml) was inoculated with a single *Agrobacterium tumefaciens* colony picked from an LB/rif plate. It was then incubated in a 27°C shaking incubator (225-250rpm) for 48 hours in the dark. Single colonies of Helper strain *E.coli* HB101/pRK2013 (kanamycin resistant) and the donor were also picked and grown up overnight at 37°C in LB liquid medium with appropriate antibiotics. Following the incubation period each bacterial culture was centrifuged at 10,000rpm for 2 minutes. The supernatants were discarded and the pellets resuspended in LB liquid medium. Aliquots of each strain (100μl each) were then mixed and spread with a sterile spreader onto an LB plate with no selection. The plate was inverted and incubated overnight at 27°C in the dark. A loopful of the overnight mating mix was then streaked onto a LB plate containing selective antibiotics (rifampicin, 100μg/ml and kanamycin 50μg/ml). The plate was inverted and incubated for 48-72 hours at 27°C in the dark. Single colonies could then be selected for use in transformation of tomato explants.

Seed sterilisation

Ailsa craig variety of tomato seeds were placed into a sterile 50ml Falcon tube. The seeds were washed with 70% ethanol for 30 seconds and the ethanol removed. 1% Virkon was then added and the tube incubated with shaking at 27°C for 20-30 minutes. 1% Virkon was then added and the tube incubated with shaking at 27°C for 20-30 minutes. The seeds were then washed with sterile dH_2O (~500ml) through a sterile sieve.

Seed sowing

MS3S medium (125ml) was poured per sterile double Magenta pot (Sigma) and allowed to set.

Five sterile seeds were then sown in each pot and incubated for 5 weeks in a control temperature room (27°C) under 5 cool white light tubes with 16 hours photoperiod and 70% relative humidity.

Explant preparation

Plates were prepared for explant preparation by the addition of MS3C5ZR medium to petri dishes (25 plates per litre of medium). A sterile 8.5cm filter disc was then placed onto each plate. Plates were wrapped in cling film and stored at room temperature. Explants were taken under aseptic conditions for 5 week old seedlings. 1-1.5cm sections from above cotyledons were cut and all leaves, roots and leaf nodes were removed. The explants were placed on a filter disc on pre-incubation medium (10 per plate as prepared in step 1. The plates were then sealed and stored at 26°C with low light intensity.

A. tumefaciens culture preparation

Several *A. tumefaciens* colonies from triparental mating containing either pVB6_TSEC-LML or pVB6_35S_TSEC-LML were picked and used to inoculate LB liquid medium (10mls) containing kanamycin (50μg/ml). The culture was incubated in a shaking incubator (225-250rpm) for 24 hours at 27°C. The overnight culture (10mls) was added to LB liquid medium (50mls) containing kanamycin (50μg/ml). This second culture was then incubated for 24 hours at 27°C in a shaking incubator (225-250rpm).

The *A. tumefaciens* culture (40mls) was then briefly centrifuged in a bench-top centrifuge (up to 3,000rpm) to remove clumps of growth. The supernatant was then carefully collected into a sterile 50ml Falcon tube. The supernatant was spun at 3,000rpm in a bench-top centrifuge for 10 minutes and the supernatant discarded. The pellet was resuspended in MS3S (30mls) by vortexing. The culture was diluted to 1/10th with MS3S and the optical density (OD) at 550nm measured with MS3S as a blank. The OD was adjusted to 0.1 with MS3S 20-25mls of culture was prepared for every 50 explants transformed.

Transformation of explants

50 explants were prepared as above (5 plates) and were transferred into petri dishes and 25ml of *A. tumefaciens* solution per petri dish poured over them. They were then incubated at room temperature for 10 minutes before being transferred to petri dishes containing a double layer of sterile filter paper. The explants were then transferred to plates

containing MS3SC5ZR medium (10 per plate). The plates were sealed and then incubated in a control temperature room (27°C) for 48 hours.

Selection

The explants were transferred to selection media MS3C5RCK (10 explants per plate) and sealed before returning to the control temperature room for 2 weeks.

Subculture of explants

Following selection explants were subcultured every 2 weeks on MS3C5ZRCK medium. When shoots developed they were carefully excised and transferred to Phytatrays (Sigma) containing MS3C5CK. DNA samples for PCR analysis were collected when shoots were sufficiently developed. Once the shoots rooted they were transferred to the glasshouse where initially they were placed in vermiculite with 1g/L Osmocote slow release fertiliser and then once roots were established they were transferred to soil.

Constructs for transformation

pVB6_35S-TSEC-LML and pVB6-TSEC-LML are shown in diagrammatic form in Figures 7 and 6 respectively.

Analysis of transformants

1. All transformation were tested for the transgene, using PCR with *E.coli* Dxps-specific primers:

Forward: 5'-GCG CCG CTA TTT ACT CGA-3'

Reverse: 5'TTT CTC TGG CGT GCC GCC-3'

2. Those that were PCR positive were tested by Southern blot analysis for the number of inserts, using the nptII probe.

3. Single insert transformation were then analysed for Dxps expression using RT-PCR, and the primers described in 1, above.

4. Expressing lines were tested for DXPS protein levels using Western blots with an antibody specific for the *E.coli* protein. A band ca. 69 kDa was found, showing both expression of transgene and cleavage of the transit peptide from the mature protein.

5. Seed was collected from all single insert lines for sowing.

6. T1 progeny were cultivated for pigment analysis and inheritance of phenotype.

Isoprenoids constitute a large group of compounds many of which are of high economic value. The condensation of (hydroxy)thiamin, derived from the decarboxylation of pyruvate, with glyceraldehyde-3-phosphate to yield 1-deoxy-D- xylulose-5-phosphate, is considered to be the first reaction in the mevalonate-independent pathway for IPP and ultimately isoprenoid biosynthesis. The data presented show that increasing the rate of DXP synthesis in *E. coli* results in increased isoprenoid biosynthesis. This finding can therefore be utilised to optimise the industrial production of isoprenoids from bacteria. The

manipulation of enzyme activities important in the biosynthesis of specific isoprenoids in concert with DXPS may be employed to bioengineer the production of specific, high value isoprenoids in *E. coli* or another suitable cell or organism such as in plants where increased isoprenoid production could be used for improving crop flavour, fragrance and colour. Alternatively, crops could be engineered to produce increased concentrations of isoprenoids with pharmaceutical and/or nutritional properties.

TABLE 1. DXP synthase activity in *E. coli homogenates*

	Specific activity nmol/min/mg protein	**Fold increase in activity**
Control	5.8 ± 0.07	1.0
B. subtilis	11.5 ± 0.58	2.0
Syn. sp. 6803	10.4 ± 0.24	1.8

References

[1] Dictionary of Natural Products on CD-ROM (v. 5.1). (1996) Chapman and Hall.

[2] McCaskill, D. and Croteau, R. (1998) TIBTECH, 16, 349-355.

[3] Horbach, S., Sahm, H. and Welle, R. (1993) FEMS Microbiol. Lett. 111: 135-140.

[4] Rohmer, M., Knani, M., Simonin, P., Sutter, B. and Sham, H. (1993) Biochem. J. 295, 517-524.

[5] Rohmer, M., Seeman, M., Horbach, S., Bringer-Meyer, S., Sahm, H. (1996) J. Am. Chem. Soc. 118, 2564-2566.

[6] Schwender, J., Seeman, M., Lichtenthaler, H. K. and Rohmer, M. (1996) Biochem. J. 316, 73-80.

[7] Eisenreich, W., Menhard, B., Hylands, P. J., Zenk, M. H., Bacher, A. (1996) Proc. Natl. Acad. Sci. USA 93, 6431-6436.

[8] Lichtenthaler, H. K., Schwender, J., Disch, A. and Rohmer, M. (1997) FEBS Lett. 400, 271-274.

[9] Eisenreich, W., Sagner, S., Zenk, M. H. and Bacher, A. (1997) Tetrahedron Lett. 38, 3889-3892.

[10] Zeidler, J. G., Lichtenthaler, H. K., May, H. U. and Lichtenthaler, F. W. Z. (1997) Naturforsch. 52c, 15-23.

[11] Lois, L. M., Campos, N., Putra, S. R., Danielsen, K., Rohmer, M. and Boronat, A. (1998) Proc. Natl. Acad. Sci. USA 95, 2105-2110.

[12] Sprenger, G. A., Sch÷rken, U., Wiegert, T., Grolle, S., de Graaf, A. A., Taylor, S. V., Begley, T. P., Bringer-Meyer, S. and Sahm, H. (1997) Proc. Natl. Acad. Sci. USA 94, 12857-12862.

[13] Lange, M. B., Wildung, M. R., McCaskill, D. and Croteau, R. (1998) Proc. Natl. Acad. Sci. USA 95, 2100-2104.

[14] Bouvier, F., d'Harlingue, A., Suire, C., Backhaus, R. A. and Camara, B. (1998) Plant Physiol. 117, 1423-1431.

[15] Mochs, C. P., Tan, L., Cunningham, F. X. and Della Penna, D. 9th International Conference on Arabidopsis Research, University of Wisconsin-Madison, USA, June 24-29, 1998 (abstr).

[16] Mandel, M. A., Feldmann, K. A., Herrera-Estrella, L., Rocha-Sosa and M., Leon, P. (1996) Plant J. 9, 649-658.

[17] Takahashi, S., Kuzuyama, T., Watanabe, H. and Seto, H. (1998) Proc. Natl. Acad. Sci. USA 95, 9879-9884.

[18] Miller, J. H. (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[19] Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M. and Stanier, R. Y. (1979) J. Gen. Microbiol. 111, 1-61.

[20] Misawa, N., Satomi, Y., Kondo, K., Yokoyama, A., Kajiwara, S., Saito, T., Ohtani, T. and Miki, W. (1995) J. Bacteriol. 177, 6575-6584.

[21] Caelles, C., Ferrer, A., Balcells, L., Hegardt, F. G and Boronat, A. (1989) Plant Mol Biol. 13, 627-638.

[22] Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular cloning : A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

[23] Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi,C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M. and Tabata, S. (1996) DNA Res. 3, 109-136.

[24] Mizuno, M., Masuda, S., Takemaru, K., Hosono, S., Sato, T., Takeuchi, M. and Kobayashi, Y. (1996) Microbiol. 142, 3103-3111.

[25] Bradford, J. M. (1976) Anal. Biochem. 72, 248-254.

[26] Zweig, G., Sherma, J. (1988) CRC Handbook of Chromatography, Vol. 1., Fat-Soluble Pigments,

CRC Press, Boca Raton, FL.

[27] Yoshida, H., Kotani, Y., Ochiai, K., Araki, K. (1998) J. Gen. Appl. Microbiol. 44, 19-26.

[28] Van Breeman, R. B. (1995) Anal. Chem. 67, 2004-2009.

[29] Collins, M. (1985) Meths. Microbiol. 18, 329-366.

[30] Hawkins, C. F., Borges, A. and Perham, R. N. (1989) FEBS Lett. 255, 77-82.

[31] Lindqvist, Y., Schneider, G., Ermler, U. and Sundstr÷m, M. (1992) EMBO J. 11, 2373-2379.

[32] Collins, M. D. and Jones, D. (1981) Microbial. Rev. 45, 316-354.

[33] Wallace, B. J. and Young, I. G. (1977) Biochim. Biophys. Acta 461, 75-83.

[34] Hemni, H., Ochnuma, S., Nagaoka, K. and Nishino, T. (1998) J. Biochem. 123, 1088-1096.

Claims

1. A method of manipulating isoprenoid expression in a plant or plant cell having a mevalonate independent isopentyl diphosphate synthesising pathway, which method comprises altering the activity of the enzyme 1-deoxy-D-xylulose-5-phosphate synthase (DXPS), or a functional equivalent thereof.

2. A method according to claim 1 wherein said isoprenoid production is increased by enhancing the activity or expression of said DXPS or lowered by inhibiting the activity or expression of said DXPS enzyme.

3. A method according to claim 2 wherein said enhanced DXPS activity occurs by transformation of said plant or plant cell with a vector comprising a nucleic acid molecule encoding said DXPS operably linked to an expression control sequence and optionally a reporter molecule

4. A method according to claim 3 wherein said DXPS encoded by said nucleic acid sequence is endogenous to said plant or plant cell.

5. A method according to claim 3 or 4 wherein said vector comprises one or more nucleic acid sequences encoding a polypeptide(s) capable of producing a desired isoprenoid.

6. A method according to claim 3 or 4 wherein said plant or plant cell is transformed with a further vector comprising one or more nucleic acid sequences encoding a polypeptide(s) capable of producing a

desired isoprenoid.

7. A method according to any of claims 3 to 6 wherein said vector comprising said nucleic acid sequence(s) encoding said DXPS and/or said polypeptide(s) capable of producing said isoprenoid further comprises a nucleic acid sequence of either a tissue specific promoter and/or encoding a plastid transit peptide.

8. A method according to any of claims 5 to 7 wherein said desired isoprenoid is one conferring a nutritional benefit.

9. A method according to claim 8 wherein said isoprenoid comprises any of the carotenoids, vitamins E, B1 or B6, chlorophylls, phenylquinones or diterpenes.

10. A plant or plant cell which has a mevalonate independent IPP biosynthetic pathway and which is transformed or transfected with a vector comprising a nucleic acid sequence encoding DXPS or a functional equivalent, derivative or bioprecursor thereof operably linked to an expression control sequence.

11. A plant or plant cell according to claim 10 wherein said vector further comprises a nucleic acid molecule encoding a reporter molecule.

12. A plant or plant cell according to claim 10 or 11 which further comprises a vector comprising one or more nucleic acid sequences encoding one or more polypeptides capable of producing a desired isoprenoid.

13. A plant or plant cell according to claim 12 wherein said desired isoprenoid comprises any of the carotenoids, vitamin E, B1 or B6, chlorophylls, phenylquinones, or diterpenes.

14. A method of manipulating isoprenoid expression in a cell or organism having a mevalonate independent isopentyl diphosphate synthesising pathway, which method comprises altering the activity of the enzyme 3-deoxy-D-xylulose-5-phosphate synthase (DXPS) or a functional equivalent thereof by transforming said cell or organism with a vector comprising a nucleic acid optionally linked to an expression control sequence and operably a reporter molecule, and a further vector comprising one or more nucleic acid sequences encoding a polypeptide(s) capable of producing a desired isoprenoid.

15. A method according to claim 14, wherein said nucleic acid sequence encoding said DXPS is endogenous to said cell or organism.

16. A method according to claim 14 or 15, wherein said cell is any of a bacterial, yeast or algal cell.

17. A method according to claim 16, wherein said bacterial cell is *E. coli*.

18. A method according to claim 14 or 15 wherein said organism is a plant.

19. A method according to any of claims 14 to 18 wherein said cell is any of a bacterial, yeast or algal cell.

20. A method according to any of claims 14 to 19

wherein said bacterial cell is *E. coli*.

21. A method according to claim 20 wherein said organism is a plant.

22. A method of identifying a compound which modulates isoprenoid activity or expression, said method comprising contacting said compound to be tested with a molecule which is a component of the mevalonate independent IPP biosynthetic pathway and which molecule undergoes a reaction catalysed by DXPS activity in the presence of an appropriate reactant, in the presence of DXPS or a functional equivalent thereof and monitoring the yield of a product of the reaction when compared to the same reaction performed in the absence of the compound to be tested.

23. A method according to claim 22 wherein said molecule comprises pyruvate and said appropriate reactant comprises glyceraldehyde-3-phosphate or vice versa.

24. A transgenic cell, tissue or organism having a mevalonate independent IPP biosynthetic pathway and increased isoprenoid activity which cell, tissue or organism comprises at least one transgene capable of expressing DXPS or a functional equivalent thereof.

25. A transgenic cell, tissue or organism according to claim 24, which comprises at least one additional copy of any of the nucleic acid sequences identified in Figure 3, or the complement thereof.

26. A transgenic cell, tissue or organism according to claim 24 or 25, further comprising a transgene capable of expressing one or more polypeptides capable of producing a desired

isoprenoid, or a functional equivalent.

27. A transgenic cell, tissue or organism according to any of claims 24 to 26, wherein said organism is a plant.

28. A transgenic cell tissue or organism according to claim 27, wherein said plant is of the Lycopersicon spp.

29. Progeny of the organism according to any of claims 24 to 28 having increased isoprenoid activity.

30. A transformed plant comprising a transgene capable of expressing DXPS from *E.coli* having the sequence according to Figure 8 and which plant comprises a higher level of isoprenoid than an untransformed plant.

31. A transformed plant according to claim 30 comprising any of constructs pVB6_TSEC_LML or pVB6_35S_TSEC-LML illustrated in Figures 6 and 7.

32. A transformed plant according to claim 30 or 31 wherein said plant is a tomato plant.

33. A tomato fruit produced by a plant according to claim 32 and having a higher level of isoprenoid activity than a wild type fruit.

34. A seed produced by a plant according to claim 32 and having a higher level of isoprenoid activity than a seed from a wild type plant.

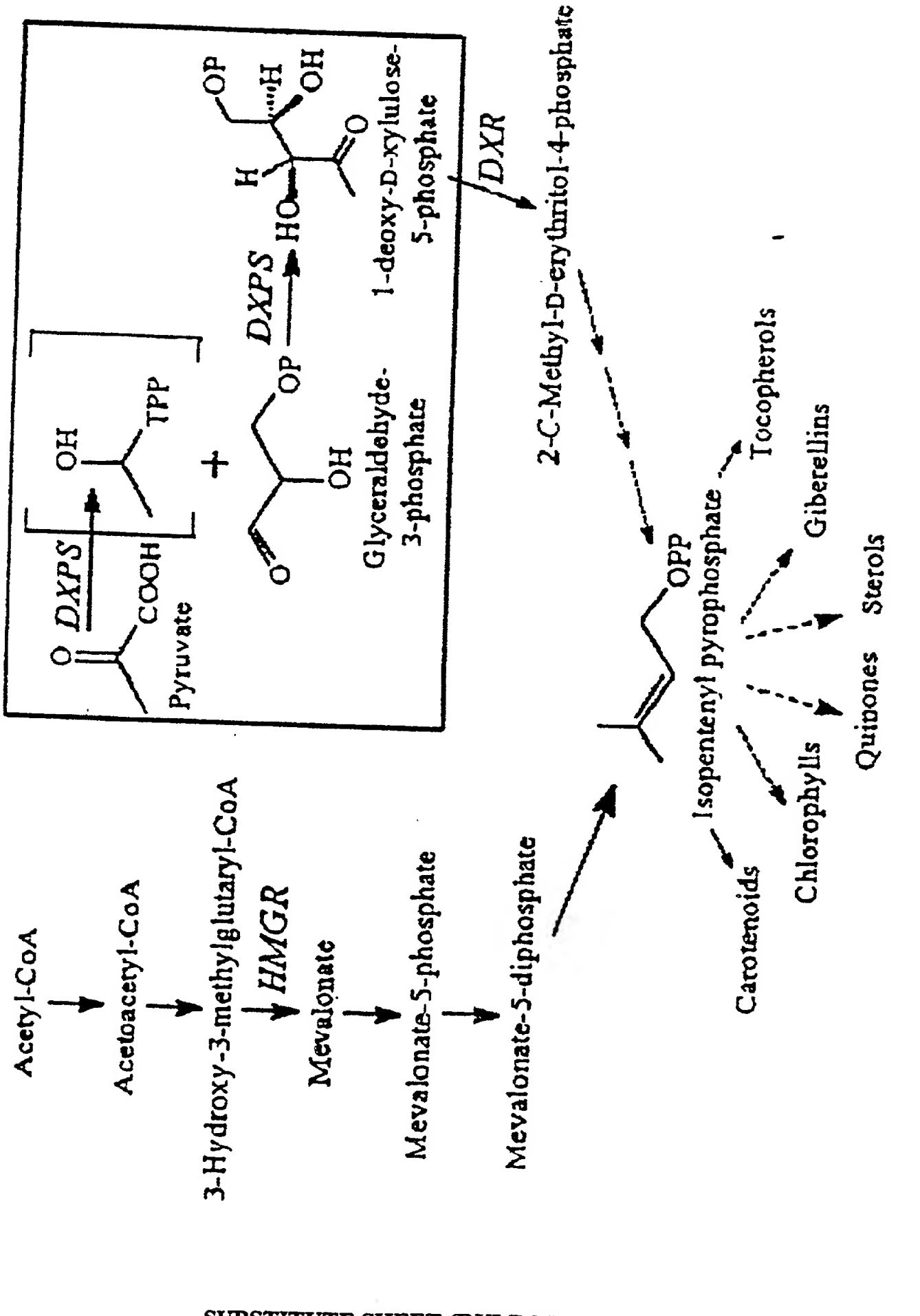
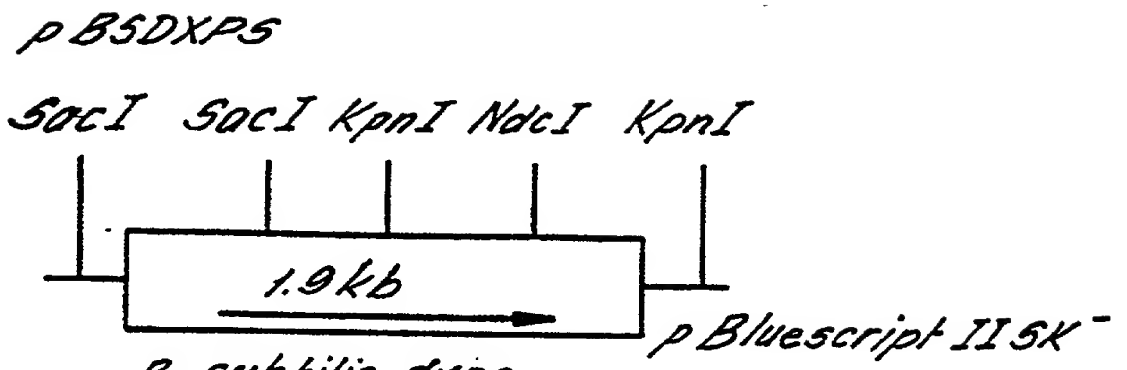
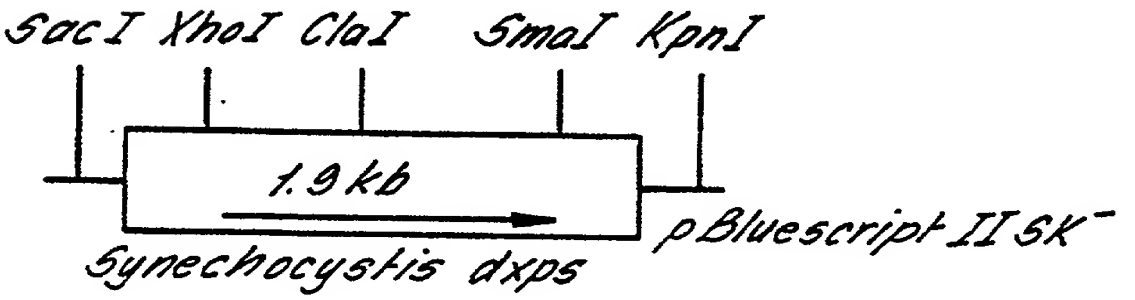

B
DXPS
TPP
OH
COOH
Pyruvate
+
DXPS
OP
HO
H
OH
O
OP
O
OH
Glyceraldehyde-3-phosphate
1-deoxy-D-xylulose-5-phosphate
DXR
2-C-Methyl-D-erythritol-4-phosphate
OPP
Isopentenyl pyrophosphate
Tocopherols
Giberellins
Sterols
Carotenoids
Chlorophylls
Quinones
A
Acetyl-CoA
Acetoacetyl-CoA
3-Hydroxy-3-methylglutaryl-CoA
HMGR
Mevalonate
Mevalonate-5-phosphate
Mevalonate-5-diphosphate

FIG. 1.

FIG. 2.


pBSDXPS
SacI SacI KpnI NdcI KpnI
1.9kb
pBluescript II SK⁻
B. subtilis dxps


SacI XhoI ClaI SmaI KpnI
1.9kb
pBluescript II SK⁻
Synechocystis dxps

```
synxxx1   ---------MDLLSIQDPSELKNMSIDELEKLSDEIRQELITSLSASGGHIGPNLGVVELT
bacillus  ---------MDLLSIQDPSFLKNMSIDELEKLSDEIRQELITSLSASGGHIGPNLGVVELT
Ecolix0   MSFDIAKYPTLALVDSTQELRLLPKESLPKLCDELRRYLLDSVSRSSGHFASGLGTVELT

synxxx1   VALHKEFNSPKDKFLWDVGHQSYVHKLLTGRGKEFATLRQYKGLCGFPKRSESEHDVWET
bacillus  VALHKEFNSPKDKFLWDVGHQSYVHKLLTGRGKEFATLRQYKGLCGFPKRSESEHDVWET
Ecolix0   VALHYVYNTPFDQLIWDVGHQAYPHKILTGRRDKIGTIRQKGGLHPFPWRGESEYDVLSV

synxxx1   GHSSTSLSGAMGMAAARDIKGTDEYIIPIIGDGALTGGMALEALNHIGDEKKDMIVILND
bacillus  GHSSTSLSGAMGMAAARDIKGTDEYIIPIIGDGALTGGMALEALNHIGDEKKDMIVILND
Ecolix0   GHSSTSISAGIGIAVAAEKEGKNRRTVCVIGDGAITAGMAFEAMNHAGDIRPDMLVILND

synxxx1   NEMSIAPNVGAIHSMLGRLRTAGKYQWVKDELEYLFKKIPAVGGKLAATAERVKDSLKYM
bacillus  NEMSIAPNVGAIHSMLGRLRTAGKYQWVKDELEYLFKKIPAVGGKLAATAERVKDSLKYM
Ecolix0   NEMSISENVGALNNHLAQLLSGKLYSSLREGGKKVFSGVPPIKELLKRTEEHIKG----M

synxxx1   LVSGMFFEELGFTYLGPVDGHSYHELIENLQYAKKTKGPVLLHVITKKGKGYKPAETDTI
bacillus  LVSGMFFEELGFTYLGPVDGHSYHELIENLQYAKKTKGPVLLHVITKKGKGYKPAETDTI
Ecolix0   VVPGTLFEELGFNYIGPVDGHDVLGLITTLKNMRDLKGPQFLHIMTKKGRGYEPAEKDPI

synxxx1   GTWHGTGPYKINTGDFVKPKAAAPSWSGLVSGTVQRMAREDGRIVAITPAMPVGSKLEGF
bacillus  GTWHGTGPYKINTGDFVKPKAAAPSWSGLVSGTVQRMAREDGRIVAITPAMPVGSKLEGF
Ecolix0   -TFHAVPKFDPSSGCLPKSSGGLPSYSKIFGDWLCETAAKDNKLMAITPAMREGSGMVEF

synxxx1   AKEFPDRMFDVGIAEQHAATMAAAMAMQGMKPFLAIYSTFLQRAYDQVVHDICRQNANVF
bacillus  AKEFPDRMFDVGIAEQHAATMAAAMAMQGMKPFLAIYSTFLQRAYDQVVHDICRQNANVF
Ecolix0   SRKFPDRYFDVAIAEQHAVTFAAGLAIGGYKPIVAIYSTFLQRAYDQVLHDVAIQKLPVL

synxxx1   IGIDRAGLVGADGETHQGVFDIAFMRHIPNMVLMMPKDENEGQHMVHTALSYDEGPIAMR
bacillus  IGIDRAGLVGADGETHQGVFDIAFMRHIPNMVLMMPKDENEGQHMVHTALSYDEGPIAMR
Ecolix0   FAIDRAGIVGADGQTHQGAFDLSYLRCIPEMVIMTPSDENECRQMLYTGYHYNDGPSAVR

synxxx1   FPRGNGLGVKMDEQLKTIPIGTWEVLRPGNDAVILTFGTTIEMAIEAAEELQKEGLSVRV
bacillus  FPRGNGLGVKMDEQLKTIPIGTWEVLRPGNDAVILTFGTTIEMAIEAAEELQKEGLSVRV
Ecolix0   YPRGNAVGVELTP-LEKLPIGKGIVKRRGEKLAILNFGTLMPEAAKVAESLN-----ATL

synxxx1   VNAREIKPIDEKMMKSILKEGLPILTIEEAVLEGGFGSSILEFAHDQGEYHTPIDRMGIP
bacillus  VNAREIKPIDEKMMKSILKEGLPILTIEEAVLEGGFGSSILEFAHDQGEYHTPIDRMGIP
Ecolix0   VDMRFVKPLDEALILEMAASHEALVTVEENAIMGGAGSGVNEVLMAHR-KPVPVLNIGLP

synxxx1   DRFIEHGSVTALLEEIGLTKQQVANRIRLLMPPKTHKGIGS
bacillus  DRFIEHGSVTALLEEIGLTKQQVANRIRLLMPPKTHKGIGS
Ecolix0   DFFIPQGTQEEMRAELGLDAAGMEAKIKAWLA----------
```

FIG. 3.

FIG. 4.

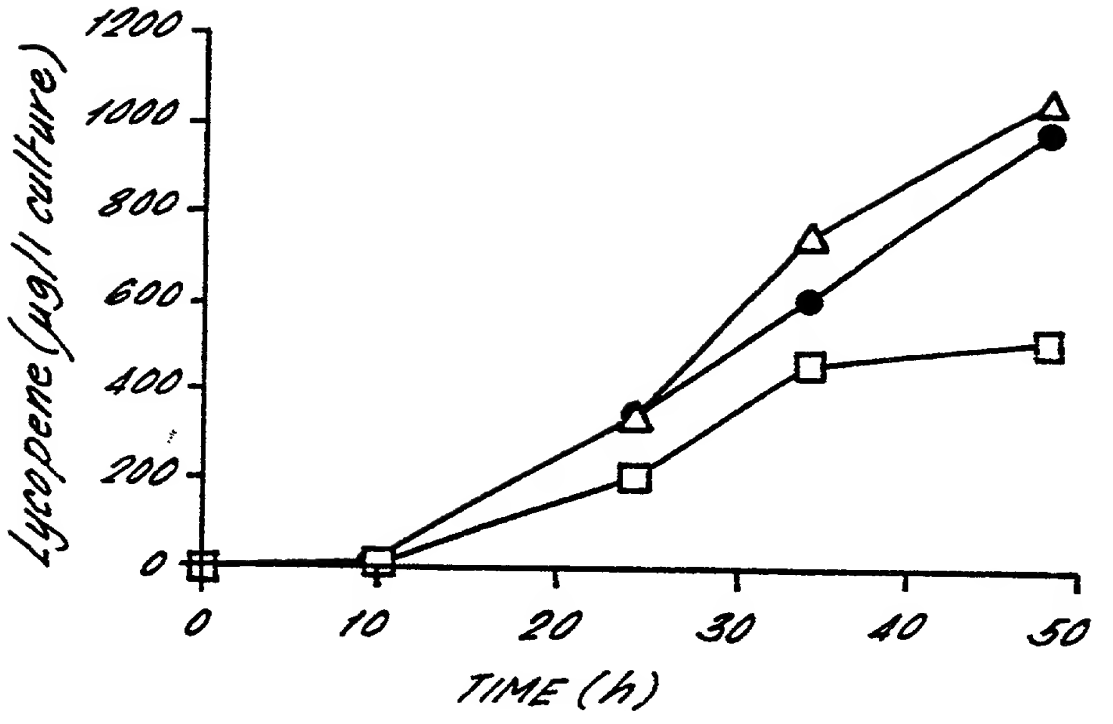
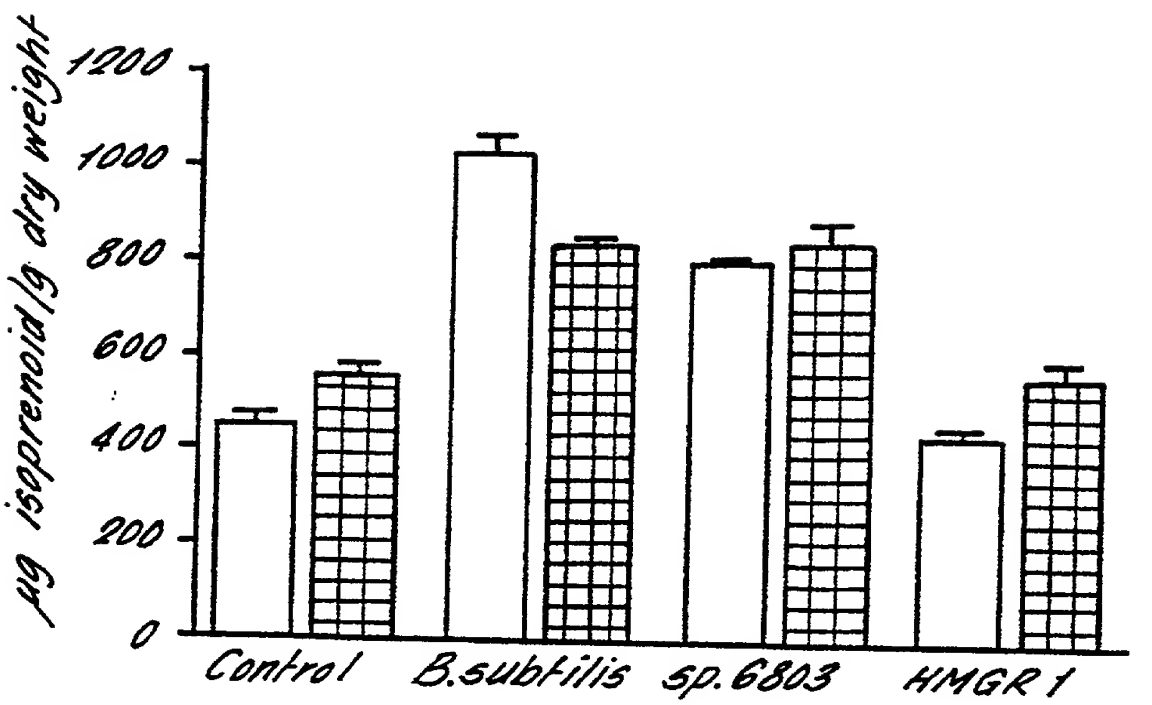

Lycopene (µg/l culture)
1200
1000
800
600
400
200
0
0
10
20
30
40
50
TIME (h)

FIG. 5.


mg isoprenoid/g dry weight
1200
1000
800
600
400
200
0
Control
B.subtilis
sp.6803
HMGR1

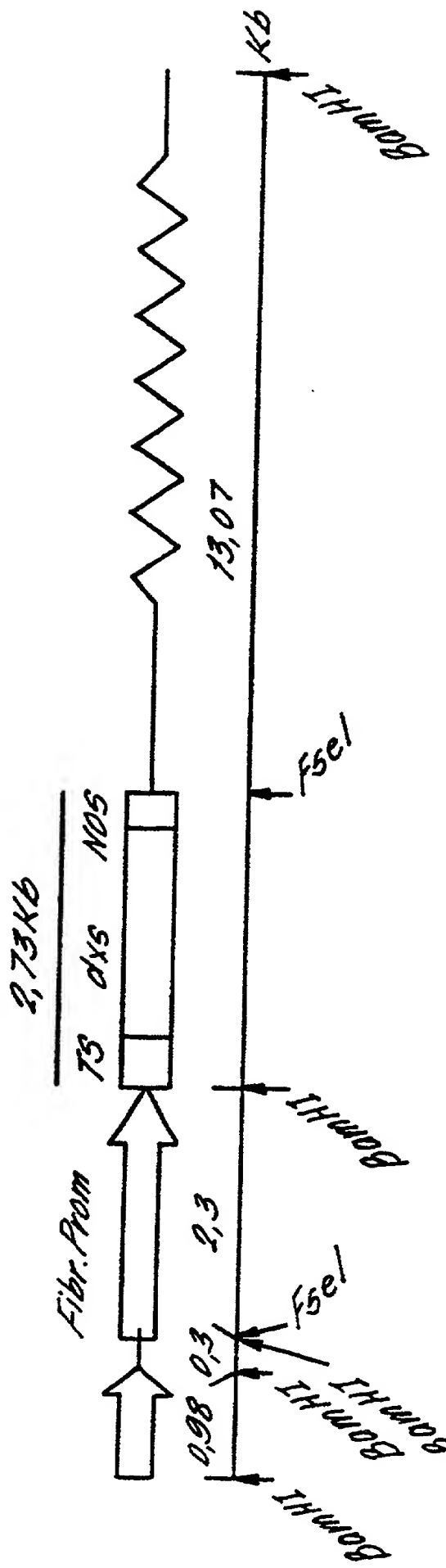
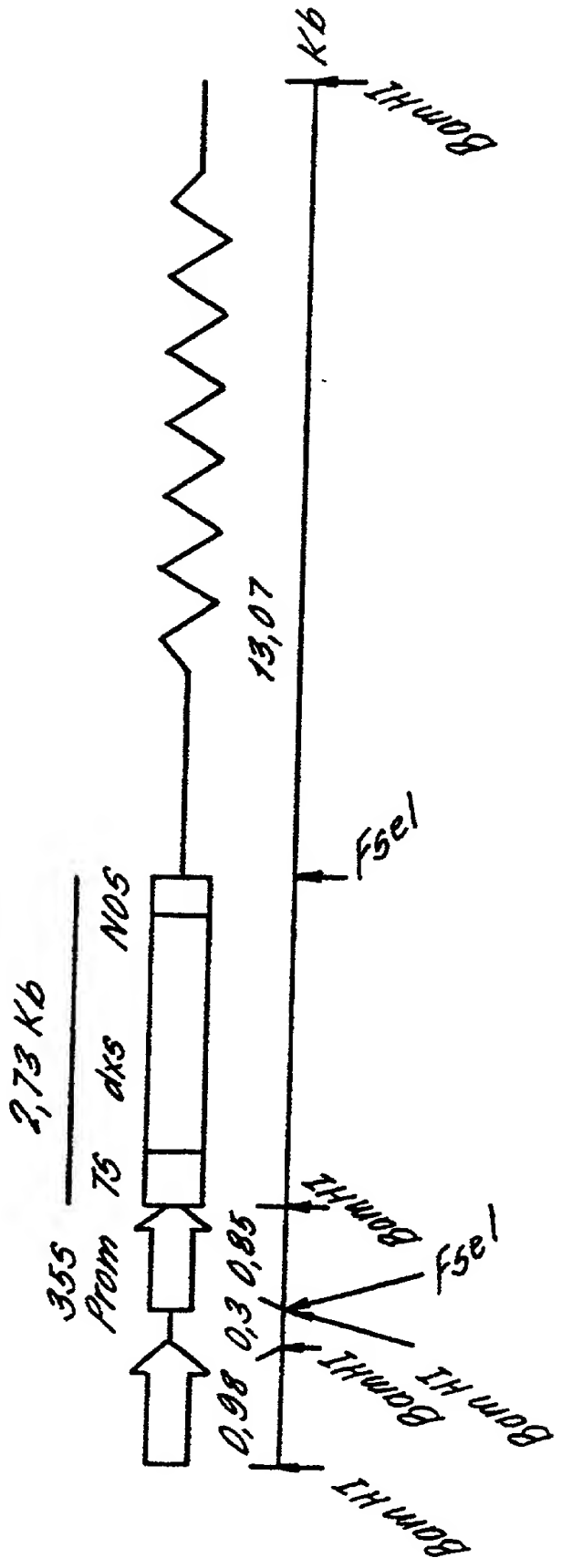


FIG. 6.

Construction name: pVB6_TSEC_LML
Fibr.Prom: Fibrilin promoter
TS: transit sequence
dXS: deoxyxylulose synthase gene from E. coli
Date: October 1998

7/8

FIG. 7.



Construction name: pVBC_355_TSEC_LML
355 Prom: 355 promoter
TS: transit sequence
dxs: deoxyxylulose synthase gene from E.coli
Date: October 1998

FIG. 8.

msfdiakyptlalvdstqelrllpkeslpklcdelrrylldsvsrssghfasglgtveltvalhyvyntpfdqliwdvghqa
yphkiltgrrdkigtirqkgglhpfpwrgeseydvlsvghsstsisagigiavaaekegknrrtvcvigdgaitagmafe
amnhagdirpdmlvilndnemsisenvgalnnhlaqllsgklysslreggkkvfsgvppikellkrteehikgmvvp
gtlfeelgfnyigpvdghdvlglittlknmrdlkgpqflhimtkkgrgyepaekdpitfhavpkfdpssgclpkssggllp
syskifgdwlcetaakdnklmaitpamregsgmvefsrkfpdryfdvaiaeqhavtfaaglaiggykpivaiystflqr
aydqvlhdvaiqklpvlfaidragivgadgqthqgafdlsylrcipemvimtpsdenecrqmlytgyhyndgpsavr
yprgnavgveltpleklpigkgivkrrgeklailnfgtlmpeaakvaeslnatlvdmrfvkpldealilemaashealvt
veenaimggagsgvnevlmahrkpvpvlniglpdffipqgtqeemraelgldaagmeakikawla

FIG. 9.

malcayafpgilnrtgvvsdsskatplfsgwihgtdlqflfqhklthevkkrsrvvqaslsesgeyytqrpptpildtvny
pihmknlslkelkqladelrsdtifnvsktgghlgsslgvveltvalhyvfnapqdrilwdvghqsyphkiltgrrdkms
tlrqtdglagftkrseseydcfg

Attorney Docket No. B0192/7031

DECLARATION FOR PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am an original, first and joint inventor of the subject matter which is claimed and for which a patent is sought on the invention entitled

MANIPULATING ISOPRENOID EXPRESSION

the specification of which is attached hereto unless the following is checked:

[X] was filed on July 27, 2001, as United States Application No. 09/890,229, bearing attorney docket no. B0192/7031, and was amended on July 27, 2001.

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, §119(a)-(d) or §365(b) of any foreign application(s) for patent or inventor's certificate, or section 365(a) of any PCT International application designating at least one country other than the United States listed below and have also identified below any foreign application for patent or inventor's certificate or PCT International application having a filing date before that of the application on which priority is claimed:

Prior Foreign PCT International Application(s)and any priority claims under 35 U.S.C. §§119 and 365(a),(b):

			Priority Claimed	
9901902.8 (Number)	**Great Britain** (Country-if PCT, so indicate)	**January 28, 1999** (DD/MM/YY Filed)	**[X]** YES	[ ] NO
(Number)	(Country-if PCT, so indicate)	(DD/MM/YY Filed)	[ ] YES	[ ] NO
(Number)	(Country-if PCT, so indicate)	(DD/MM/YY Filed)	[ ] YES	[ ] NO

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below:

(Application Number)	(filing date)
(Application Number)	(filing date)

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s), or §365(c) of any PCT International application(s) designating the United States of America listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application:

(Application No.)	(filing date)	(status-patented, pending, abandoned)
(Application No.)	(filing date)	(status-patented, pending, abandoned)

PCT International Applications designating the United States:

(PCT Appl. No.)	(U.S. Ser. No.)	(PCT filing date)	(status-patented,pending,abandoned)

I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith:

Robert M. Abrahamsen	40,886	Jason M. Honeyman	31,624	Edward J. Russavage	43,069
Konstantinos Andrikopoulos	P-48,915	Robert E. Hunt	39,231	Stanley Sacks	19,900
Eric Amundsen	46,518	Ronald J. Kransdorf	20,004	Christopher S. Schultz	37,929
John N. Anastasi	37,765	Peter C. Lando	34,654	Robert A. Skrivanek, Jr.	41,316
Ilan Barzilay	46,540	M. Brad Lawrence	47,210	Alan W. Steele	45,128
Carole Boelitz	P-48,958	Helen C. Lockhart	39,248	Mark Steinberg	40,829
Gary S. Engelson	35,128	Matthew B. Lowrie	38,228	Joseph Teja, Jr.	45,157
Neil P. Ferraro	39,188	William R. McClellan	29,409	Maryanne Trevisan	P-48,207
Thomas G. Field III	45,596	Daniel P. McLoughlin	46,066	John R. Van Amsterdam	40,212
Stephen R. Finch	42,534	James H. Morris	34,681	Robert H. Walat	46,324
Edward R. Gates	31,616	Timothy J. Oyer	36,628	Kristin D. Wheeler	43,583
Richard F. Giunta	36,149	Edward F. Perlman	28,105	Lisa E. Winsor	44,405
Lawrence M. Green	29,384	Elizabeth R. Plumer	36,637	David Wolf	17,528
George L. Greenfield	17,756	Michael J. Pomianek	46,190	Douglas R. Wolf	36,971
James M. Hanifin, Jr.	39,213	Randy J. Pritzker	35,986		
Steven J. Henry	27,900				

Address all telephone calls to John R. Van Amsterdam at telephone no. (617) 720-3500. Address all correspondence to:

John R. Van Amsterdam
c/o Wolf, Greenfield & Sacks, P.C.,
Federal Reserve Plaza
600 Atlantic Avenue
Boston, MA 02210-2211

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment,

or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Peter M Bramley 16/10/07

Inventor's signature **Date**

Full name of first inventor: **Peter M. Bramley**
Citizenship: Great Britain
Residence: Egham, Surrey, UK GB
Post Office Address: Royal Holloway University of London, School of Biological Sciences, Department of Biochemistry, Egham, Surrey TW20 0EX, United Kingdom

22/10/07.

Inventor's signature **Date**

Full name of second inventor: **Mark Harker**
Citizenship: Great Britain
Residence: St. Neots, UK GB
Post Office Address: 10 Langwood Close, Eaton Sorcon, St. Neots PE19 3QN, United Kingdom

SEQUENCE LISTING

<110> Bramley, Peter Michael
Harker, Mark

<120> Manipulating Isoprenoid Expression

<130> B0192/7031

<140> PCT/GB00/00263
<141> 2000-01-28

<150> GB 9901902.8
<151> 1999-01-28

<160> 12

<170> PatentIn version 3.0

<210> 1
<211> 640
<212> PRT
<213> Synechocystis sp.

<400> 1

Met His Ile Ser Glu Leu Thr His Pro Asn Glu Leu Lys Gly Leu Ser
1 5 10 15

Ile Arg Glu Leu Glu Glu Val Ser Arg Gln Ile Arg Glu Lys His Leu
20 25 30

Gln Thr Val Ala Thr Ser Gly Gly His Leu Gly Pro Gly Leu Gly Val
35 40 45

Val Glu Leu Thr Val Ala Leu Tyr Ser Thr Leu Asp Leu Asp Lys Asp
50 55 60

Arg Val Ile Trp Asp Val Gly His Gln Ala Tyr Pro His Lys Met Leu
65 70 75 80

Thr Gly Arg Tyr His Asp Phe His Thr Leu Arg Gln Lys Asp Gly Val
85 90 95

Ala Gly Tyr Leu Lys Arg Ser Glu Ser Arg Phe Asp His Phe Gly Ala
100 105 110

Gly His Ala Ser Thr Ser Ile Ser Ala Gly Leu Gly Met Ala Leu Ala
115 120 125

Arg Asp Ala Lys Gly Glu Asp Phe Lys Val Val Ser Ile Ile Gly Asp
130 135 140

Gly Ala Leu Thr Gly Gly Met Ala Leu Glu Ala Ile Asn His Ala Gly
145 150 155 160

His Leu Pro His Thr Arg Leu Met Val Ile Leu Asn Asp Asn Glu Met
165 170 175

```
Ser Ile Ser Pro Asn Val Gly Ala Ile Ser Arg Tyr Leu Asn Lys Val
            180                 185                 190

Arg Leu Ser Ser Pro Met Gln Phe Leu Thr Asp Asn Leu Glu Glu Gln
        195                 200                 205

Ile Lys His Leu Pro Phe Val Gly Asp Ser Leu Thr Pro Glu Met Glu
    210                 215                 220

Arg Val Lys Glu Gly Met Lys Arg Leu Val Val Pro Lys Val Gly Ala
225                 230                 235                 240

Val Ile Glu Glu Leu Gly Phe Lys Tyr Phe Gly Pro Ile Asp Gly His
                245                 250                 255

Ser Leu Gln Glu Leu Ile Asp Thr Phe Lys Gln Ala Glu Lys Val Pro
            260                 265                 270

Gly Pro Val Phe Val His Val Ser Thr Thr Lys Gly Lys Gly Tyr Asp
        275                 280                 285

Leu Ala Glu Lys Asp Gln Val Gly Tyr His Ala Gln Ser Pro Phe Asn
    290                 295                 300

Leu Ser Thr Gly Lys Ala Tyr Pro Ser Ser Lys Pro Lys Pro Pro Ser
305                 310                 315                 320

Tyr Ser Lys Val Phe Ala His Thr Leu Thr Thr Leu Ala Lys Glu Asn
                325                 330                 335

Pro Asn Ile Val Gly Ile Thr Ala Ala Met Ala Thr Gly Thr Gly Leu
            340                 345                 350

Asp Lys Leu Gln Ala Lys Leu Pro Lys Gln Tyr Val Asp Val Gly Ile
        355                 360                 365

Ala Glu Gln His Ala Val Thr Leu Ala Ala Gly Met Ala Cys Glu Gly
    370                 375                 380

Ile Arg Pro Val Val Ala Ile Tyr Ser Thr Phe Leu Gln Arg Gly Tyr
385                 390                 395                 400

Asp Gln Ile Ile His Asp Val Cys Ile Gln Lys Leu Pro Val Phe Phe
                405                 410                 415

Cys Leu Asp Arg Ala Gly Ile Val Gly Ala Asp Gly Pro Thr His Gln
            420                 425                 430

Gly Met Tyr Asp Ile Ala Tyr Leu Arg Cys Ile Pro Asn Leu Val Leu
        435                 440                 445

Met Ala Pro Lys Asp Glu Ala Glu Leu Gln Gln Met Leu Val Thr Gly
    450                 455                 460

Val Asn Tyr Thr Gly Gly Ala Ile Ala Met Arg Tyr Pro Arg Gly Asn
465                 470                 475                 480

Gly Ile Gly Val Pro Leu Met Glu Glu Gly Trp Glu Pro Leu Glu Ile
                485                 490                 495
```

```
Gly Lys Ala Glu Ile Leu Arg Ser Gly Asp Asp Val Leu Leu Leu Gly
            500                 505                 510

Tyr Gly Ser Met Val Tyr Pro Ala Leu Gln Thr Ala Glu Leu Leu His
        515                 520                 525

Glu His Gly Ile Glu Ala Thr Val Val Asn Ala Arg Phe Val Lys Pro
    530                 535                 540

Leu Asp Thr Glu Leu Ile Leu Pro Leu Ala Glu Arg Ile Gly Lys Val
545                 550                 555                 560

Val Thr Met Glu Glu Gly Cys Leu Met Gly Gly Phe Gly Ser Ala Val
                565                 570                 575

Ala Glu Ala Leu Met Asp Asn Asn Val Leu Val Pro Leu Lys Arg Leu
            580                 585                 590

Gly Val Pro Asp Ile Leu Val Asp His Ala Thr Pro Glu Gln Ser Thr
        595                 600                 605

Val Asp Leu Gly Leu Thr Pro Ala Gln Met Ala Gln Asn Ile Met Ala
    610                 615                 620

Ser Leu Phe Lys Thr Glu Thr Glu Ser Val Val Ala Pro Gly Val Ser
625                 630                 635                 640
```

<210> 2
<211> 633
<212> PRT
<213> Bacillus subtilis

<400> 2

```
Met Asp Leu Leu Ser Ile Gln Asp Pro Ser Phe Leu Lys Asn Met Ser
1               5                   10                  15

Ile Asp Glu Leu Glu Lys Leu Ser Asp Glu Ile Arg Gln Phe Leu Ile
            20                  25                  30

Thr Ser Leu Ser Ala Ser Gly Gly His Ile Gly Pro Asn Leu Gly Val
        35                  40                  45

Val Glu Leu Thr Val Ala Leu His Lys Glu Phe Asn Ser Pro Lys Asp
    50                  55                  60

Lys Phe Leu Trp Asp Val Gly His Gln Ser Tyr Val His Lys Leu Leu
65                  70                  75                  80

Thr Gly Arg Gly Lys Glu Phe Ala Thr Leu Arg Gln Tyr Lys Gly Leu
                85                  90                  95

Cys Gly Phe Pro Lys Arg Ser Glu Ser Glu His Asp Val Trp Glu Thr
            100                 105                 110

Gly His Ser Ser Thr Ser Leu Ser Gly Ala Met Gly Met Ala Ala Ala
        115                 120                 125
```

```
Arg Asp Ile Lys Gly Thr Asp Glu Tyr Ile Ile Pro Ile Ile Gly Asp
    130                 135                 140

Gly Ala Leu Thr Gly Gly Met Ala Leu Glu Ala Leu Asn His Ile Gly
145                 150                 155                 160

Asp Glu Lys Lys Asp Met Ile Val Ile Leu Asn Asp Asn Glu Met Ser
                165                 170                 175

Ile Ala Pro Asn Val Gly Ala Ile His Ser Met Leu Gly Arg Leu Arg
            180                 185                 190

Thr Ala Gly Lys Tyr Gln Trp Val Lys Asp Glu Leu Glu Tyr Leu Phe
        195                 200                 205

Lys Lys Ile Pro Ala Val Gly Gly Lys Leu Ala Ala Thr Ala Glu Arg
    210                 215                 220

Val Lys Asp Ser Leu Lys Tyr Met Leu Val Ser Gly Met Phe Phe Glu
225                 230                 235                 240

Glu Leu Gly Phe Thr Tyr Leu Gly Pro Val Asp Gly His Ser Tyr His
                245                 250                 255

Glu Leu Ile Glu Asn Leu Gln Tyr Ala Lys Lys Thr Lys Gly Pro Val
            260                 265                 270

Leu Leu His Val Ile Thr Lys Lys Gly Lys Gly Tyr Lys Pro Ala Glu
        275                 280                 285

Thr Asp Thr Ile Gly Thr Trp His Gly Thr Gly Pro Tyr Lys Ile Asn
    290                 295                 300

Thr Gly Asp Phe Val Lys Pro Lys Ala Ala Ala Pro Ser Trp Ser Gly
305                 310                 315                 320

Leu Val Ser Gly Thr Val Gln Arg Met Ala Arg Glu Asp Gly Arg Ile
                325                 330                 335

Val Ala Ile Thr Pro Ala Met Pro Val Gly Ser Lys Leu Glu Gly Phe
            340                 345                 350

Ala Lys Glu Phe Pro Asp Arg Met Phe Asp Val Gly Ile Ala Glu Gln
        355                 360                 365

His Ala Ala Thr Met Ala Ala Ala Met Ala Met Gln Gly Met Lys Pro
    370                 375                 380

Phe Leu Ala Ile Tyr Ser Thr Phe Leu Gln Arg Ala Tyr Asp Gln Val
385                 390                 395                 400

Val His Asp Ile Cys Arg Gln Asn Ala Asn Val Phe Ile Gly Ile Asp
                405                 410                 415

Arg Ala Gly Leu Val Gly Ala Asp Gly Glu Thr His Gln Gly Val Phe
            420                 425                 430

Asp Ile Ala Phe Met Arg His Ile Pro Asn Met Val Leu Met Met Pro
        435                 440                 445
```

Lys Asp Glu Asn Glu Gly Gln His Met Val His Thr Ala Leu Ser Tyr
450 455 460

Asp Glu Gly Pro Ile Ala Met Arg Phe Pro Arg Gly Asn Gly Leu Gly
465 470 475 480

Val Lys Met Asp Glu Gln Leu Lys Thr Ile Pro Ile Gly Thr Trp Glu
485 490 495

Val Leu Arg Pro Gly Asn Asp Ala Val Ile Leu Thr Phe Gly Thr Thr
500 505 510

Ile Glu Met Ala Ile Glu Ala Ala Glu Glu Leu Gln Lys Glu Gly Leu
515 520 525

Ser Val Arg Val Val Asn Ala Arg Phe Ile Lys Pro Ile Asp Glu Lys
530 535 540

Met Met Lys Ser Ile Leu Lys Glu Gly Leu Pro Ile Leu Thr Ile Glu
545 550 555 560

Glu Ala Val Leu Glu Gly Gly Phe Gly Ser Ser Ile Leu Glu Phe Ala
565 570 575

His Asp Gln Gly Glu Tyr His Thr Pro Ile Asp Arg Met Gly Ile Pro
580 585 590

Asp Arg Phe Ile Glu His Gly Ser Val Thr Ala Leu Leu Glu Glu Ile
595 600 605

Gly Leu Thr Lys Gln Gln Val Ala Asn Arg Ile Arg Leu Leu Met Pro
610 615 620

Pro Lys Thr His Lys Gly Ile Gly Ser
625 630

<210> 3
<211> 620
<212> PRT
<213> Escherichia coli

<400> 3

Met Ser Phe Asp Ile Ala Lys Tyr Pro Thr Leu Ala Leu Val Asp Ser
1 5 10 15

Thr Gln Glu Leu Arg Leu Leu Pro Lys Glu Ser Leu Pro Lys Leu Cys
20 25 30

Asp Glu Leu Arg Arg Tyr Leu Leu Asp Ser Val Ser Arg Ser Ser Gly
35 40 45

His Phe Ala Ser Gly Leu Gly Thr Val Glu Leu Thr Val Ala Leu His
50 55 60

Tyr Val Tyr Asn Thr Pro Phe Asp Gln Leu Ile Trp Asp Val Gly His
65 70 75 80

Gln Ala Tyr Pro His Lys Ile Leu Thr Gly Arg Arg Asp Lys Ile Gly

```
                85                  90                  95

Thr Ile Arg Gln Lys Gly Gly Leu His Pro Phe Pro Trp Arg Gly Glu
            100                 105                 110

Ser Glu Tyr Asp Val Leu Ser Val Gly His Ser Ser Thr Ser Ile Ser
        115                 120                 125

Ala Gly Ile Gly Ile Ala Val Ala Ala Glu Lys Glu Gly Lys Asn Arg
    130                 135                 140

Arg Thr Val Cys Val Ile Gly Asp Gly Ala Ile Thr Ala Gly Met Ala
145                 150                 155                 160

Phe Glu Ala Met Asn His Ala Gly Asp Ile Arg Pro Asp Met Leu Val
                165                 170                 175

Ile Leu Asn Asp Asn Glu Met Ser Ile Ser Glu Asn Val Gly Ala Leu
            180                 185                 190

Asn Asn His Leu Ala Gln Leu Leu Ser Gly Lys Leu Tyr Ser Ser Leu
        195                 200                 205

Arg Glu Gly Gly Lys Lys Val Phe Ser Gly Val Pro Pro Ile Lys Glu
    210                 215                 220

Leu Leu Lys Arg Thr Glu Glu His Ile Lys Gly Met Val Val Pro Gly
225                 230                 235                 240

Thr Leu Phe Glu Glu Leu Gly Phe Asn Tyr Ile Gly Pro Val Asp Gly
                245                 250                 255

His Asp Val Leu Gly Leu Ile Thr Thr Leu Lys Asn Met Arg Asp Leu
            260                 265                 270

Lys Gly Pro Gln Phe Leu His Ile Met Thr Lys Lys Gly Arg Gly Tyr
        275                 280                 285

Glu Pro Ala Glu Lys Asp Pro Ile Thr Phe His Ala Val Pro Lys Phe
    290                 295                 300

Asp Pro Ser Ser Gly Cys Leu Pro Lys Ser Ser Gly Gly Leu Pro Ser
305                 310                 315                 320

Tyr Ser Lys Ile Phe Gly Asp Trp Leu Cys Glu Thr Ala Ala Lys Asp
                325                 330                 335

Asn Lys Leu Met Ala Ile Thr Pro Ala Met Arg Glu Gly Ser Gly Met
            340                 345                 350

Val Glu Phe Ser Arg Lys Phe Pro Asp Arg Tyr Phe Asp Val Ala Ile
        355                 360                 365

Ala Glu Gln His Ala Val Thr Phe Ala Ala Gly Leu Ala Ile Gly Gly
    370                 375                 380

Tyr Lys Pro Ile Val Ala Ile Tyr Ser Thr Phe Leu Gln Arg Ala Tyr
385                 390                 395                 400

Asp Gln Val Leu His Asp Val Ala Ile Gln Lys Leu Pro Val Leu Phe
```

405 410 415

Ala Ile Asp Arg Ala Gly Ile Val Gly Ala Asp Gly Gln Thr His Gln
420 425 430

Gly Ala Phe Asp Leu Ser Tyr Leu Arg Cys Ile Pro Glu Met Val Ile
435 440 445

Met Thr Pro Ser Asp Glu Asn Glu Cys Arg Gln Met Leu Tyr Thr Gly
450 455 460

Tyr His Tyr Asn Asp Gly Pro Ser Ala Val Arg Tyr Pro Arg Gly Asn
465 470 475 480

Ala Val Gly Val Glu Leu Thr Pro Leu Glu Lys Leu Pro Ile Gly Lys
485 490 495

Gly Ile Val Lys Arg Arg Gly Glu Lys Leu Ala Ile Leu Asn Phe Gly
500 505 510

Thr Leu Met Pro Glu Ala Ala Lys Val Ala Glu Ser Leu Asn Ala Thr
515 520 525

Leu Val Asp Met Arg Phe Val Lys Pro Leu Asp Glu Ala Leu Ile Leu
530 535 540

Glu Met Ala Ala Ser His Glu Ala Leu Val Thr Val Glu Glu Asn Ala
545 550 555 560

Ile Met Gly Gly Ala Gly Ser Gly Val Asn Glu Val Leu Met Ala His
565 570 575

Arg Lys Pro Val Pro Val Leu Asn Ile Gly Leu Pro Asp Phe Phe Ile
580 585 590

Pro Gln Gly Thr Gln Glu Glu Met Arg Ala Glu Leu Gly Leu Asp Ala
595 600 605

Ala Gly Met Glu Ala Lys Ile Lys Ala Trp Leu Ala
610 615 620

<210> 4
<211> 184
<212> PRT
<213> Lycopersicon esculentum

<400> 4

Met Ala Leu Cys Ala Tyr Ala Phe Pro Gly Ile Leu Asn Arg Thr Gly
1 5 10 15

Val Val Ser Asp Ser Ser Lys Ala Thr Pro Leu Phe Ser Gly Trp Ile
20 25 30

His Gly Thr Asp Leu Gln Phe Leu Phe Gln His Lys Leu Thr His Glu
35 40 45

Val Lys Lys Arg Ser Arg Val Val Gln Ala Ser Leu Ser Glu Ser Gly
50 55 60

```
Glu Tyr Tyr Thr Gln Arg Pro Pro Thr Pro Ile Leu Asp Thr Val Asn
65                  70                  75                  80

Tyr Pro Ile His Met Lys Asn Leu Ser Leu Lys Glu Leu Lys Gln Leu
                85                  90                  95

Ala Asp Glu Leu Arg Ser Asp Thr Ile Phe Asn Val Ser Lys Thr Gly
            100                 105                 110

Gly His Leu Gly Ser Ser Leu Gly Val Val Glu Leu Thr Val Ala Leu
        115                 120                 125

His Tyr Val Phe Asn Ala Pro Gln Asp Arg Ile Leu Trp Asp Val Gly
    130                 135                 140

His Gln Ser Tyr Pro His Lys Ile Leu Thr Gly Arg Arg Asp Lys Met
145                 150                 155                 160

Ser Thr Leu Arg Gln Thr Asp Gly Leu Ala Gly Phe Thr Lys Arg Ser
                165                 170                 175

Glu Ser Glu Tyr Asp Cys Phe Gly
            180
```

<210> 5
<211> 16032
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: plasmid pVB6_35S_TSEC_LML

<220>
<221> unsure
<222> (1641)..(1641)
<223> n = a, c, g, or t/u

<220>
<221> unsure
<222> (1644)..(1644)
<223> n = a, c, g, or t/u

<220>
<221> unsure
<222> (1652)..(1652)
<223> n = a, c, g, or t/u

<220>
<221> unsure
<222> (1656)..(1656)
<223> n = a, c, g, or t/u

<400> 5
atgattgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggctattc 60

ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca 120

gcgcaggggc gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg 180

cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcagctgtg ctcgacgttg 240

tcactgaagc gggaagggac tggctgctat tgggcgaagt gccggggcag gatctcctgt 300

catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg cggcggctgc 360

atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc atcgagcgag 420

cacgtactcg gatggaagcc ggtcttgtcg atcaggatga tctggacgaa gagcatcagg 480

ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgcg catgcccgac ggcgaggatc 540

tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat ggccgctttt 600

ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac atagcgttgg 660

ctacccgtga tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc ctcgtgcttt 720

acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt gacgagttct 780

tctgagaatt ctttagcctt taaattgcta gttttcgtta aatggtcata acttgagcta 840

tggacctcca aattaaattt cgggcatacg ctcaaatccc aaattacgat acggagctac 900

cggaactgtc aaaatactga tccgggtccg tttgctaaaa acgttgacca aagtccacta 960

agttgagttt taaaacttta tttcacattt taatccattt tttacatgaa aactttccgg 1020

aaaatacgga gtatgcacgc aagtcgagga atgataaatg gtacttttcg aagttttaga 1080

actcaaaatt acttattaaa tttaaagatg acattttggg tcatcacatt gatgaaaatt 1140

ttgacattaa tatctgagaa ctttctttga ccttttttcga ttctaatcca atcaattcaa 1200

cagtgtaagg tgaagcagtc aatttaaagg aaggccttta aattctaaaa tattgtactt 1260

ttcctgcgct tctaaaagtg aacgacaaag aaaaaatagt tattcttgaa cttaatattg 1320

tacaatagga taaattttaa ctatctataa aaagagaaca aaaccttaat ctcttcaaaa 1380

taatattata agaagtaaca taattgtcaa atgaaataca cataagaagc acataaattt 1440

aaatgccgta ttaaacttac agtatactat agcggaagtt ggcttgataa aggaacgctg 1500

aggagagtag ccgatggtga aacactaaca tcaagtgcaa aagaaagaaa aactgaaaac 1560

agaagatgaa tgtttgaagt gggtaaaaga ttacttaaaa gataggtttg gttaacaaat 1620

gattgtgact gttacycrsc nscntmdstr anstsncatg gctttgtgtg cttatgcatt 1680

tcctgggatt ttgaacagga ctggtgtggt ttcagattct tctaaggcaa ccccttttgtt 1740

ctctggatgg attcatggaa cagatctgca gtttttgttc caacacaagc ttactcatga 1800

ggtcaagaaa aggtcacgtg tggttcaggc ttccttatca gaatctggag aatactacac 1860

acagagaccg ccaacgccta ttttggacac tgtgaactat cccattcata tgaaaaatct 1920

gtctctgaag gaacttaaac aactagcaga tgaactaagg tcagatacaa ttttcaatgt 1980

atcaaagact gggggtcacc ttggctcaag tcttggtgtt gttgagctga ctgttgctct 2040

tcattatgtc ttcaatgcac cgcaagatag gattctctgg gatgttggtc atcagtctta 2100

tcctcacaaa atcttgactg gtagaaggga caagatgtcg acattaaggc agacagatgg 2160

tcttgcagga tttactaagc gatcggagag tgaatatgat tgctttgatg agttttgata 2220

ttgccaaata cccgaccctg gcactggtcg actccaccca ggagttacga ctgttgccga 2280

aagagagttt accgaaactc tgcgacgaac tgcgccgcta tttactcgac agcgtgagcc 2340

gttccagcgg gcacttcgcc tccgggctgg gcacggtcga actgaccgtg gcgctgcact 2400

atgtctacaa caccccgttt gaccaattga tttgggatgt ggggcatcag gcttatccgc 2460

ataaaatttt gaccggacgc cgcgacaaaa tcggcaccat ccgtcagaaa ggcggtctgc 2520

acccgttccc gtggcgcggc gaaagcgaat atgacgtatt aagcgtcggg cattcatcaa 2580

cctccatcag tgccggaatt ggtattgcgg ttgctgccga aaaagaaggc aaaaatcgcc 2640

gcaccgtctg tgtcattggc gatggcgcga ttaccgcagg catggcgttt gaagcgatga 2700

atcacgcggg cgatatccgt cctgatatgc tggtgattct caacgacaat gaaatgtcga 2760

tttccgaaaa tgtcggcgcg ctcaacaacc atctggcaca gctgctttcc ggtaagcttt 2820

actcttcact gcgcgaaggc gggaaaaaag ttttctctgg cgtgccgcca attaaagagc 2880

tgctcaaacg caccgaagaa catattaaag gcatggtagt gcctggcacg ttgtttgaag 2940

agctgggctt taactacatc ggcccggtgg acggtcacga tgtgctgggg cttatcacca 3000

cgctaaagaa catgcgcgac ctgaaaggcc cgcagttcct gcatatcatg accaaaaaag 3060

gtcgtggtta tgaaccggca gaaaaagacc cgatcacttt ccacgccgtg cctaaatttg 3120

atccctccag cggttgtttg ccgaaaagta gcggcggttt gccgagctat tcaaaaatct 3180

ttggcgactg gttgtgcgaa acggcagcga aagacaacaa gctgatggcg attactccgg 3240

cgatgcgtga aggttccggc atggtcgagt tttcacgtaa attcccggat cgctacttcg 3300

acgtggcaat tgccgagcaa cacgcggtga cctttgctgc gggtctggcg attggtgggt 3360

acaaacccat tgtcgcgatt tactccactt tcctgcaacg cgcctatgat caggtgctgc 3420

atgacgtggc gattcaaaag cttccggtcc tgttcgccat cgaccgcgcg ggcattgttg 3480

gtgctgacgg tcaaacccat cagggtgctt ttgatctctc ttacctgcgc tgcataccgg 3540

aaatggtcat tatgaccccg agcgatgaaa acgaatgtcg ccagatgctc tataccggct 3600

atcactataa cgatggcccg tcagcggtgc gctacccgcg tggcaacgcg gtcggcgtgg 3660

aactgacgcc gctggaaaaa ctaccaattg gcaaaggcat tgtgaagcgt cgtggcgaga 3720

aactggcgat ccttaacttt ggtacgctga tgccagaagc ggcgaaagtc gccgaatcgc 3780

tgaacgccac gctggtcgat atgcgttttg tgaaaccgct tgatgaagcg ttaattctgg 3840

aaatggccgc cagccatgaa gcgctggtca ccgtagaaga aaacgccatt atgggcggcg 3900

caggcagcgg cgtgaacgaa gtgctgatgg cccatcgtaa accagtaccc gtgctgaaca 3960

ttggcctgcc ggacttcttt attccgcaag gaactcagga agaaatgcgc gccgaactcg 4020

gcctcgatgc cgctggtatg gaagccaaaa tcaaggcctg gctggcataa cgaattcccg 4080

atctagtaac atagatgaca ccgcgcgcga taatttatcc tagtttgcgc gctatatttt 4140

gttttctatc gcgtattaaa tgtataattg cgggactcta atcataaaaa cccatctcat 4200

aaataacgtc atgcacctga atagatcttg gacaagcgtt aggcctatct gtgcattaca 4260

tgttaattat tacatgctta acgtaattca acagaaatta tatgataatc atcgcaagac 4320

cggcaacagg attcaatctt aagaaacttt attgccaaat gtttgaacga tcggggaaat 4380

tcgagctccc gggctggttg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa 4440

cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgccggc gttgtggata 4500

cctcgcggaa aacttggccc tcactgacag atgaggggcg gacgttgaca cttgaggggc 4560

cgactcaccc ggcgcggcgt tgacagatga ggggcaggct cgatttcggc cggcgacgtg 4620

gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgcgagtt tcccacagat 4680

gatgtggaca agcctgggga taagtgccct gcggtattga cacttgaggg gcgcgactac 4740

tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc 4800

gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt 4860

ccgcccgttt ttcggccacc gctaacctgt cttttaacct gcttttaaac caatatttat 4920

aaaccttgtt tttaaccagg gctgcgccct gtgcgcgtga ccgcgcacgc cgaaggggggg 4980

tgccccccct tctcgaaccc tcccggcccg ctaacgcggg cctcccatcc ccccaggggc 5040

tgcgcccctc ggccgcgaac ggcctcaccc caaaaatggc agcgctggca gtccttgcca 5100

ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca 5160

ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg 5220

gcggcctggg tggcggcctg cccttcactt cggccgtcgg ggcattcacg gacttcatgg 5280

cggggccggc aatttttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg 5340

tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag 5400

gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag 5460

ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 5520

tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 5580

ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 5640

tgcatggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 5700

attgggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccgatgactt 5760

tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 5820

gctgcctcag attcaggtta tgccgctcaa ttcgctgcgt atatcgcttg ctgattacgt 5880

gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca 5940

cgtcaaaggg tgaagcaggc tcataagacg ccccagcgtc gccatagtgc gttcaccgaa 6000

tacgtgcgca acaaccgtct tccggagact gtcatacgcg taaaacagcc agcgctggcg 6060

cgatttagcc ccgacatagc cccactgttc gtccatttcc gcgcagacga tgacgtcact 6120

gcccggctgt atgcgcgagg ttaccgactg cggcctgagt tttttaagtg acgtaaaatc 6180

gtgttgaggc caacgcccat aatgcgggct gttgcccggc atccaacgcc attcatggcc 6240

atatcaatga ttttctggtg cgtaccgggt tgagaagcgg tgtaagtgaa ctgcagttgc 6300

catgttttac ggcagtgaga gcagagatag cgctgatgtc cggcggtgct tttgccgtta 6360

cgcaccaccc cgtcagtagc tgaacaggag ggacagctga tagaaacaga agccactgga 6420

gcacctcaaa aacaccatca tacactaaat cagtaagttg gcagcatcac ccataattgt 6480

ggtttcaaaa tcggctccgt cgatactatg ttatacgcca actttgaaaa caactttgaa 6540

aaagctgttt tctggtattt aaggttttag aatgcaagga acagtgaatt ggagttcgtc 6600

ttgttataat tagcttcttg gggtatcttt aaatactgta gaaaagagga aggaaataat 6660

aaatggctaa aatgagaata tcaccggaat tgaaaaaact gatcgaaaaa taccgctgcg 6720

taaaagatac ggaaggaatg tctcctgcta aggtatataa gctggtggga gaaaatgaaa 6780

acctatattt aaaaatgacg gacagccggt ataaagggac cacctatgat gtggaacggg 6840

aaaaggacat gatgctatgg ctggaaggaa agctgcctgt tccaaaggtc ctgcactttg 6900

aacggcatga tggctggagc aatctgctca tgagtgaggc cgatggcgtc ctttgctcgg 6960

aagagtatga agatgaacaa agccctgaaa agattatcga gctgtatgcg gagtgcatca 7020

ggctctttca ctccatcgac atatcggatt gtccctatac gaatagctta gacagccgct 7080

tagccgaatt ggattactta ctgaataacg atctggccga tgtggattgc gaaaactggg 7140

aagaagacac tccatttaaa gatccgcgcg agctgtatga ttttttaaag acggaaaagc 7200

ccgaagagga acttgtcttt tcccacggcg acctgggaga cagcaacatc tttgtgaaag 7260

atggcaaagt aagtggcttt attgatcttg ggagaagcgg cagggcggac aagtggtatg 7320

acattgcctt ctgcgtccgg tcgatcaggg aggatatcgg ggaagaacag tatgtcgagc 7380

tatttttga cttactgggg atcaagcctg attgggagaa aataaaatat tatattttac 7440

tggatgaatt gttttagtac ctagatgtgg cgcaacgatg ccggcgacaa gcaggagcgc 7500

accgacttct tccgcatcaa gtgttttggc tctcaggccg aggcccacgg caagtatttg 7560

ggcaaggggt cgctggtatt cgtgcagggc aagattcgga ataccaagta cgagaaggac 7620

ggccagacgg tctacgggac cgacttcatt gccgataagg tggattatct ggacaccaag 7680

gcaccaggcg ggtcaaatca ggaataaggg cacattgccc cggcgtgagt cggggcaatc 7740

ccgcaaggag ggtgaatgaa tcggacgttt gaccggaagg catacaggca agaactgatc 7800

gacgcggggt tttccgccga ggatgccgaa accatcgcaa gccgcaccgt catgcgtgcg 7860

ccccgcgaaa ccttccagtc cgtcggctcg atggtccagc aagctacggc caagatcgag 7920

cgcgacagcg tgcaactggc tccccctgcc ctgcccgcgc catcggccgc cgtggagcgt 7980

tcgcgtcgtc tcgaacagga ggcggcaggt ttggcgaagt cgatgaccat cgacacgcga 8040

ggaactatga cgaccaagaa gcgaaaaacc gccggcgagg acctggcaaa acaggtcagc 8100

gaggccaagc aggccgcgtt gctgaaacac acgaagcagc agatcaagga aatgcagctt 8160

tccttgttcg atattgcgcc gtggccggac acgatgcgag cgatgccaaa cgacacggcc 8220

cgctctgccc tgttcaccac gcgcaacaag aaaatcccgc gcgaggcgct gcaaaacaag 8280

gtcattttcc acgtcaacaa ggacgtgaag atcacctaca ccggcgtcga gctgcgggcc 8340

gacgatgacg aactggtgtg gcagcaggtg ttggagtacg cgaagcgcac ccctatcggc 8400

gagccgatca ccttcacgtt ctacgagctt tgccaggacc tgggctggtc gatcaatggc 8460

cggtattaca cgaaggccga ggaatgcctg tcgcgcctac aggcgacggc gatgggcttc 8520

acgtccgacc gcgttgggca cctggaatcg gtgtcgctgc tgcaccgctt ccgcgtcctg 8580

gaccgtggca agaaaacgtc ccgttgccag gtcctgatcg acgaggaaat cgtcgtgctg 8640

tttgctggcg accactacac gaaattcata tgggagaagt accgcaagct gtcgccgacg 8700

gcccgacgga tgttcgacta tttcagctcg caccgggagc cgtacccgct caagctggaa 8760

accttccgcc tcatgtgcgg atcggattcc acccgcgtga agaagtggcg cgagcaggtc 8820

ggcgaagcct gcgaagagtt gcgaggcagc ggcctggtgg aacacgcctg ggtcaatgat 8880

gacctggtgc attgcaaacg ctagggcctt gtggggtcag ttccggctgg gggttcagca 8940

gccagcgctt tactggcatt tcaggaacaa gcgggcactg ctcgacgcac ttgcttcgct 9000

cagtatcgct cgggacgcac ggcgcgctct acgaactgcc gataaacaga ggattaaaat 9060

tgacaattgt gattaaggct cagattcgac ggcttggagc ggccgacgtg caggatttcc 9120

gcgagatccg attgtcggcc ctgaagaaag ctccagagat gttcgggtcc gtttacgagc 9180

acgaggagaa aaagcccatg gaggcgttcg ctgaacggtt gcgagatgcc gtggcattcg 9240

gcgcctacat cgacggcgag atcattgggc tgtcggtctt caaacaggag gacggcccca 9300

aggacgctca caaggcgcat ctgtccggcg ttttcgtgga gcccgaacag cgaggccgag 9360

gggtcgccgg tatgctgctg cgggcgttgc cggcgggttt attgctcgtg atgatcgtcc 9420

gacagattcc aacgggaatc tggtggatgc gcatcttcat cctcggcgca cttaatattt 9480

cgctattctg gagcttgttg tttatttcgg tctaccgcct gccgggcggg gtcgcggcga 9540

cggtaggcgc tgtgcagccg ctgatggtcg tgttcatctc tgccgctctg ctaggtagcc 9600

cgatacgatt gatggcggtc ctgggggcta tttgcggaac tgcgggcgtg gcgctgttgg 9660

tgttgacacc aaacgcagcg ctagatcctg tcggcgtcgc agcgggcctg gcgggggcgg 9720

tttccatggc gttcggaacc gtgctgaccc gcaagtggca acctcccgtg cctctgctca 9780

cctttaccgc ctggcaactg gcggccggag gacttctgct cgttccagta gctttagtgt 9840

ttgatccgcc aatcccgatg cctacaggaa ccaatgttct cggcctggcg tggctcggcc 9900

tgatcggagc gggtttaacc tacttccttt ggttccgggg gatctcgcga ctcgaaccta 9960

cagttgtttc cttactgggc tttctcagcc ccagatctgg ggtcgatcag ccggggatgc 10020

atcaggccga cagtcggaac ttcgggtccc cgacctgtac cattcggtga gcaatggata 10080

ggggagttga tatcgtcaac gttcacttct aaagaaatag cgccactcag cttcctcagc 10140

ggctttatcc agcgatttcc tattatgtcg gcatagttct caagatcgac agcctgtcac 10200

ggttaagcga gaaatgaata agaaggctga taattcggat ctctgcgagg gagatgatat 10260

ttgatcacag gcagcaacgc tctgtcatcg ttacaatcaa catgctaccc tccgcgagat 10320

catccgtgtt tcaaacccgg cagcttagtt gccgttcttc cgaatagcat cggtaacatg 10380

agcaaagtct gccgccttac aacggctctc ccgctgacgc cgtcccggac tgatgggctg 10440

cctgtatcga gtggtgattt tgtgccgagc tgccggtcgg ggagctgttg gctggctggt 10500

ggcaggatat attgtggtgt aaacaaattg acgcttagac aacttaataa cacattgcgg 10560

acgtttttaa tgtactgggg tggtttttct tttcaccagt gagacgggca acagctgatt 10620

gcccttcacc gcctggccct gagagagttg cagcaagcgg tccacgctgg tttgccccag 10680

caggcgaaaa tcctgtttga tggtggttcc gaaatcggca aaatccctta taaatcaaaa 10740

gaatagcccg agatagggtt gagtgttgtt ccagtttgga acaagagtcc actattaaag 10800

aacgtggact ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt 10860

gaaccatcac ccaaatcaag ttttttgggg tcgaggtgcc gtaaagcact aaatcggaac 10920

cctaaaggga gcccccgatt tagagcttga cggggaaagc cggcgaacgt ggcgagaaag 10980

gaagggaaga aagcgaaagg agcgggcgcc attcaggctg cgcaactgtt gggaagggcg 11040

atcggtgcgg gcctcttcgc tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg 11100

attaagttgg gtaacgccag ggttttccca gtcacgacgt tgtaaaacga cggccagtga 11160

attgaagctt ggcgcgccaa gcttggtacc ctcaggacta gaccatcgtg gttaaatgat 11220

caagtgccta cttggcagaa tttctttcga gcagcctcct cctacaagtt gcatttgttg 11280

cgcttacgat aattgtcaaa gaagtaggta aaataaagac atgatctcta atattaagga 11340

taagattaaa aataagtcca ggattacccg gtcggcccat caattacttg ctgacctttg 11400

ttgccgtccc acgacttcca ttttctaacc gtccattttc atttgttttt agctatattt 11460

aatattaatg ggatataaat tataaacatt cctcctccca aaaaaataag tttaagtaat 11520

actgcaatag acagtgtttt aagccatgta attcagtaaa agttcttttt tattctgaac 11580

ctagccctaa aaaggccatg cgggtaatta gttcagtcaa ctgaatatac aacgttttga 11640

accaaagtta acatgtacag gccaatagaa gttatttgac cgtaagctta gtctctacat 11700

tcattcaacg ttcttgaatc aaagtgacct gtacaggcca atagaagtta cctgaccgta 11760

agcttagtct ctacattcat tcctctgaga cgatattcta gaagcctgct ttcaagtcta 11820

aaaggcacaa tcttttctcc tcaccacttg ttgaggtact tatgatttta aagatgaaac 11880

atttttttta cttttcccct ttaatttctt tgatttttt ttttcttggt agttggaagt 11940

acttttcata ccctagaaaa tccactgttg atctttgaaa tatcagcaat ctttgaaata 12000

atatcagcaa ccacgacacc taccattctc aaattcactc tataaaaggg taaacctttg 12060

cttacctcta tgctcactca caaggagaac aaacactcat cggtgctaca taaccgcggc 12120

tgcaggtcga cggatctgta cccggggatc cgtcgatcgt ttcgcatgat tgaacaagat 12180

ggattgcacg caggttctcc ggccgcttgg gtggagaggc tattcggcta tgactgggca 12240

caacagacaa tcggctgctc tgatgccgcc gtgttccggc tgtcagcgca ggggcgcccg 12300

gttcttttttg tcaagaccga cctgtccggt gccctgaatg aactgcagga cgaggcagcg 12360

cggctatcgt ggctggccac gacgggcgtt ccttgcgcag ctgtgctcga cgttgtcact 12420

gaagcgggaa gggactggct gctattgggc gaagtgccgg ggcaggatct cctgtcatct 12480

caccttgctc ctgccgagaa agtatccatc atggctgatg caatgcggcg gctgcatacg 12540

cttgatccgg ctacctgccc attcgaccac caagcgaaac atcgcatcga gcgagcacgt 12600

actcggatgg aagccggtct tgtcgatcag gatgatctgg acgaagagca tcaggggctc 12660

gcgccagccg aactgttcgc caggctcaag gcgcgcatgc ccgacggcga ggatctcgtc 12720

gtgacccatg gcgatgcctg cttgccgaat atcatggtgg aaaatggccg cttttctgga 12780

ttcatcgact gtggccggct gggtgtggcg gaccgctatc aggacatagc gttggctacc 12840

cgtgatattg ctgaagagct tggcggcgaa tgggctgacc gcttcctcgt gctttacggt 12900

atcgccgctc ccgattcgca gcgcatcgcc ttctatcgcc ttcttgacga gttcttctga 12960

gcgggactct ggggttcgaa atgaccgacc aagcgacgcc caacctgcca tcacgagatt 13020

tcgattccac cgccgccttc tatgaaaggt tgggcttcgg aatcgttttc cgggacgccg 13080

gctggatgat cctccagcgc ggggatctca tgctggagtt cttcgcccac ccccggatcc 13140

tctagagtcg acctgcaggc gttcaaacat ttggcaataa agtttcttaa gattgaatcc 13200

tgttgccggt cttgcgatga ttatcatata atttctgttg aattacgtta agcatgtaat 13260

aattaacatg taatgcatga cgttatttat gagatgggtt tttatgatta gagtcccgca 13320

attatacatt taatacgcga tagaaaacaa aatatagcgc gcaaactagg ataaattatc 13380

gcgcgcggtg tcatctatgt tactagatcg ggaattgcca agctggcgcg ccctgcagac 13440

cggtggatcc ggccggccga tccggtcatc ggcgggggtc ataacgtgac tcccttaatt 13500

ctccgctcat gatcagattg tcgtttcccg ccttcagttt aaactatcag tgtttgacag 13560

gatatattgg cgggtaaacc taagagaaaa gagcgtttat tagaataatc ggatatttaa 13620

aagggcgtga aaaggtttat ccgttcgtcc atttgtatgt gcatgccaac cacagggttc 13680

cccagatctg gcgccggcca gcgagacgag caagattggc cgccgcccga aacgatccga 13740

cagcgcgccc agcacaggtg cgcaggcaaa ttgcaccaac gcatacagcg ccagcagaat 13800

gccatagtgg gcggtgacgt cgttcgagtg aaccagatcg cgcaggaggc ccggcagcac 13860

cggcataatc aggccgatgc cgacagcgtc gagcgcgaca gtgctcagaa ttacgatcag 13920

gggtatgttg ggtttcacgt ctggcctccg gaccagcctc cgctggtccg attgaacgcg 13980

cggattcttt atcactgata agttggtgga catattatgt ttatcagtga taaagtgtca 14040

agcatgacaa agttgcagcc gaatacagtg atccgtgccg ccctggacct gttgaacgag 14100

gtcggcgtag acggtctgac gacacgcaaa ctggcggaac ggttgggggt tcagcagccg 14160

gcgctttact ggcacttcag gaacaagcgg gcgctgctcg acgcactggc cgaagccatg 14220

ctggcggaga atcatacgca ttcggtgccg agagccgacg acgactggcg ctcatttctg 14280

atcgggaatg cccgcagctt caggcaggcg ctgctcgcct accgcgatgg cgcgcgcatc 14340

catgccggca cgcgaccggg cgcaccgcag atggaaacgg ccgacgcgca gcttcgcttc 14400

ctctgcgagg cgggtttttc ggccggggac gccgtcaatg cgctgatgac aatcagctac 14460

ttcactgttg gggccgtgct tgaggagcag gccggcgaca gcgatgccgg cgagcgcggc 14520

ggcaccgttg aacaggctcc gctctcgccg ctgttgcggg ccgcgataga cgccttcgac 14580

gaagccggtc cggacgcagc gttcgagcag ggactcgcgg tgattgtcga tggattggcg 14640

aaaaggaggc tcgttgtcag gaacgttgaa ggaccgagaa agggtgacga ttgatcagga 14700

ccgctgccgg agcgcaaccc actcactaca gcagagccat gtagacaaca tcccctcccc 14760

ctttccaccg cgtcagacgc ccgtagcagc ccgctacggg ctttttcatg ccctgcccta 14820

gcgtccaagc ctcacggccg cgctcggcct ctctggcggc cttctggcgc tcttccgctt 14880

cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagctcact 14940

caaaggcggt aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag 15000

caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata 15060

ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc 15120

cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg 15180

ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcgc 15240

ttttccgctg cataaccctg cttcggggtc attatagcga ttttttcggt atatccatcc 15300

tttttcgcac gatatacagg attttgccaa agggttcgtg tagactttcc ttggtgtatc 15360

caacggcgtc agccgggcag gataggtgaa gtaggcccac ccgcgagcgg gtgttccttc 15420

ttcactgtcc cttattcgca cctggcggtg ctcaacggga atcctgctct gcgaggctgg 15480

ccggctaccg ccggcgtaac agatgagggc aagcggatgg ctgatgaaac caagccaacc 15540

aggaagggca gcccacctat caaggtgtac tgccttccag acgaacgaag agcgattgag 15600

gaaaaggcgg cggcggccgg catgagcctg tcggcctacc tgctggccgt cggccagggc 15660

tacaaaatca cgggcgtcgt ggactatgag cacgtccgcg agctggcccg catcaatggc 15720

gacctgggcc gcctgggcgg cctgctgaaa ctctggctca ccgacgaccc gcgcacggcg 15780

cggttcggtg atgccacgat cctcgccctg ctggcgaaga tcgaagagaa gcaggacgag 15840

cttggcaagg tcatgatggg cgtggtccgc ccgagggcag agccatgact tttttagccg 15900

ctaaaacggc cgggggggtgc gcgtgattgc caagcacgtc cccatgcgct ccatcaagaa 15960

gagcgacttc gcggagctgg tgaagtacat caccgacgag caaggcaaga ccgagcgcct 16020

ttccgacgct ca 16032

<210> 6
<211> 17460
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: plasmid

pVB6_TSEC_LML

<220>
<221> unsure
<222> (1135)..(1135)
<223> n = a, c, g, or t/u

<220>
<221> unsure
<222> (1145)..(1145)
<223> n = a, c, g, or t/u

<400> 6
atgattgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggctattc 60

ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca 120

gcgcaggggc gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg 180

cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcagctgtg ctcgacgttg 240

tcactgaagc gggaagggac tggctgctat tgggcgaagt gccggggcag gatctcctgt 300

catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg cggcggctgc 360

atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc atcgagcgag 420

cacgtactcg gatggaagcc ggtcttgtcg atcaggatga tctggacgaa gagcatcagg 480

ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgcg catgcccgac ggcgaggatc 540

tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat ggccgctttt 600

ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac atagcgttgg 660

ctacccgtga tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc ctcgtgcttt 720

acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt gacgagttct 780

tctgaaattc atgtaaagca aaaattaaaa tatgatcatc caaattattt ttatagccaa 840

tttaggtaat aatcattttt attaccactc aaattaactt cttgttaaat gaacatacat 900

atgaaaaatc tagctcgaaa tgaaacccca agaaaatctg tagtcgctat caatggtgcg 960

cactgaataa accccttact gtgcagtagc ctacaaccca catattgtga ttattgatac 1020

tacataaaat actcttaatt gaagtagtac gccatgcaat catattctca ttttggagga 1080

atccaatcct catcaaagta cccaactagc ctttctccac aaggctttga acatnattcc 1140

aatanggaat acaatcatat tctcattttg gaggagcgat tttggaaatt aatatcacat 1200

gtttaatggc ctaaatgatg gcaatacgaa cactaaatcc tttcatatat tcacttaaaa 1260

aaaagtcata taattgcttt taggagagag ataaatgatc gtattgactt agagtatact 1320

ggttaaaatt attatatatc taataattaa tttgatgttt tagtaataaa ctaccatact 1380

tgttatttt catttaaac tttactgtga aagacagaat gcacttcaga aaactaaata 1440

tctattacac atcatctgaa atggtcgtga gtgaaccttt ttgaatatat aatttaagtt 1500

gcaaatagag ttatttgaag aaaaaaaaaa aggaaaaaga aacagtataa tggctccttg 1560

tcctcttagc cctccaaaaa aggaacaaat attaatttaa ttagaaaaaa acaaccttga 1620

gcgggtcgaa ttagtaggtc tttccacatt caaattacta attttaaaaa ttattgacta 1680

cttcaagatt ctttaatctt ttttcctttg tattgagaga gtcatcaaaa agctcactcg 1740

ggacgtgaaa aaaaatttga cttgctaatt atattctata tttactaggt aaaatcctaa 1800

tcttattatc acattatcac tgggattttt cggcattatt ctagctaaat atcttgtgca 1860

attcatgtcc tccaccacaa aaaaaatgcc taattttaca actttttttgg tatgaaaaga 1920

gagtaagaaa aaggaataag agaagtctga ttgaaagaaa taaggaagga tgaataaaaa 1980

acaaagaaaa ttctactaga atttgtatgc cgttggatgt gaatggtttc caatttttt 2040

gtctctcttg tttgttcaat tttaaatttc cgccaaacag acacaaaatg atccttaact 2100

ccgctttaca agcggatagt tacgtgtctt cctctctctc ccttgttgac gtatcttaaa 2160

aactccaaac taccctgga tttttctaat ctttaattaa ggattatata tatatatata 2220

tatatatata tatatatata tatatatata tatattatga taattaataa ttaaattatt 2280

tgcacattta aaagtctatt tggattgact tattttagat gttttaaagt taaaataact 2340

tttaaataat tttagtattc gaataaacta agaaaggtgc ttataagcac tttatgcctt 2400

tacactacaa tgtaaaaatt aagtcaaaag ttattaaacg aacttattag tcaaaagtta 2460

aagcgattcc acacaagcgc tacgcttaat atttttttag ccaaagcgaa tccaaacaag 2520

cgctacgtca atattctttt ccttttcttt aattgaacca tcaatcctag atctcacttt 2580

ctctgacatg ggacctaaca ttaacaatat gagttgtggt gtgataattc gagattcctt 2640

caaccctaat aggaagtgct atagtgaaaa atgagccctc caatatccgc attcaaattt 2700

aatccgagtt taatgtagat acaatgcata atgtgggaac cgatgggaaa gaacaaaagt 2760

aataagatat acaatatgtt ttcacacata gtgagatcta gaaagggtag atagtatgcc 2820

gttggtatct ctatcttaga ggtaaagtag aaaagttgtt tccaatcgat ccgaactcaa 2880

gagaaaaaca gtttagccga tgggaaataa agaaaagaaa aactaattta ggagggtata 2940

tatgtctttg cacaaaggct aacctaaagc cctggctcgt ataaaacgcg atcattacag 3000

gattgcaaca taaacactca ttcaatcgaa gcttccctgt taatattatc attttgttcc 3060

ttctttattc actcttaaac catggatggc tttgtgtgct tatgcatttc ctgggatttt 3120

gaacaggact ggtgtggttt cagattcttc taaggcaacc cctttgttct ctggatggat 3180

tcatggaaca gatctgcagt ttttgttcca acacaagctt actcatgagg tcaagaaaag 3240

gtcacgtgtg gttcaggctt ccttatcaga atctggagaa tactacacac agagaccgcc 3300

aacgcctatt ttggacactg tgaactatcc cattcatatg aaaaatctgt ctctgaagga 3360

acttaaacaa ctagcagatg aactaaggtc agatacaatt ttcaatgtat caaagactgg 3420

gggtcacctt ggctcaagtc ttggtgttgt tgagctgact gttgctcttc attatgtctt 3480

caatgcaccg caagatagga ttctctggga tgttggtcat cagtcttatc ctcacaaaat 3540

cttgactggt agaagggaca agatgtcgac attaaggcag acagatggtc ttgcaggatt 3600

tactaagcga tcggagagtg aatatgattg ctttgatgag ttttgatatt gccaaatacc 3660

cgaccctggc actggtcgac tccacccagg agttacgact gttgccgaaa gagagtttac 3720

cgaaactctg cgacgaactg cgccgctatt tactcgacag cgtgagccgt tccagcgggc 3780

acttcgcctc cgggctgggc acggtcgaac tgaccgtggc gctgcactat gtctacaaca 3840

ccccgtttga ccaattgatt tgggatgtgg ggcatcaggc ttatccgcat aaaattttga 3900

ccggacgccg cgacaaaatc ggcaccatcc gtcagaaagg cggtctgcac ccgttcccgt 3960

ggcgcggcga aagcgaatat gacgtattaa gcgtcgggca ttcatcaacc tccatcagtg 4020

ccggaattgg tattgcggtt gctgccgaaa aagaaggcaa aaatcgccgc accgtctgtg 4080

tcattggcga tggcgcgatt accgcaggca tggcgtttga agcgatgaat cacgcgggcg 4140

atatccgtcc tgatatgctg gtgattctca acgacaatga aatgtcgatt tccgaaaatg 4200

tcggcgcgct caacaaccat ctggcacagc tgctttccgg taagctttac tcttcactgc 4260

gcgaaggcgg gaaaaaagtt ttctctggcg tgccgccaat taaagagctg ctcaaacgca 4320

ccgaagaaca tattaaaggc atggtagtgc ctggcacgtt gtttgaagag ctgggcttta 4380

actacatcgg cccggtggac ggtcacgatg tgctggggct tatcaccacg ctaaagaaca 4440

tgcgcgacct gaaaggcccg cagttcctgc atatcatgac caaaaaaggt cgtggttatg 4500

aaccggcaga aaaagacccg atcactttcc acgccgtgcc taaatttgat ccctccagcg 4560

gttgtttgcc gaaaagtagc ggcggtttgc cgagctattc aaaaatcttt ggcgactggt 4620

tgtgcgaaac ggcagcgaaa gacaacaagc tgatggcgat tactccggcg atgcgtgaag 4680

gttccggcat ggtcgagttt tcacgtaaat tcccggatcg ctacttcgac gtggcaattg 4740

ccgagcaaca cgcggtgacc tttgctgcgg gtctggcgat tggtgggtac aaacccattg 4800

tcgcgattta ctccactttc ctgcaacgcg cctatgatca ggtgctgcat gacgtggcga 4860

ttcaaaagct tccggtcctg ttcgccatcg accgcgcggg cattgttggt gctgacggtc 4920

aaacccatca gggtgctttt gatctctctt acctgcgctg cataccggaa atggtcatta 4980

tgaccccgag cgatgaaaac gaatgtcgcc agatgctcta taccggctat cactataacg 5040

atggcccgtc agcggtgcgc tacccgcgtg gcaacgcggt cggcgtggaa ctgacgccgc 5100

tggaaaaact accaattggc aaaggcattg tgaagcgtcg tggcgagaaa ctggcgatcc 5160

ttaactttgg tacgctgatg ccagaagcgg cgaaagtcgc cgaatcgctg aacgccacgc 5220

tggtcgatat gcgttttgtg aaaccgcttg atgaagcgtt aattctggaa atggccgcca 5280

gccatgaagc gctggtcacc gtagaagaaa acgccattat gggcggcgca ggcagcggcg 5340

tgaacgaagt gctgatggcc catcgtaaac cagtacccgt gctgaacatt ggcctgccgg 5400

acttctttat tccgcaagga actcaggaag aaatgcgcgc cgaactcggc ctcgatgccg 5460

ctggtatgga agccaaaatc aaggcctggc tggcataacg aattcccgat ctagtaacat 5520

agatgacacc gcgcgcgata atttatccta gtttgcgcgc tatattttgt tttctatcgc 5580

gtattaaatg tataattgcg ggactctaat cataaaaacc catctcataa ataacgtcat 5640

gcacctgaat agatcttgga caagcgttag gcctatctgt gcattacatg ttaattatta 5700

catgcttaac gtaattcaac agaaattata tgataatcat cgcaagaccg gcaacaggat 5760

tcaatcttaa gaaactttat tgccaaatgt ttgaacgatc ggggaaattc gagctcccgg 5820

gctggttgcc ctcgccgctg ggctggcggc cgtctatggc cctgcaaacg cgccagaaac 5880

gccgtcgaag ccgtgtgcga gacaccgcgg ccgccggcgt tgtggatacc tcgcggaaaa 5940

cttggccctc actgacagat gaggggcgga cgttgacact tgaggggccg actcacccgg 6000

cgcggcgttg acagatgagg ggcaggctcg atttcggccg gcgacgtgga gctggccagc 6060

ctcgcaaatc ggcgaaaacg cctgatttta cgcgagtttc ccacagatga tgtggacaag 6120

cctggggata agtgccctgc ggtattgaca cttgaggggc gcgactactg acagatgagg 6180

ggcgcgatcc ttgacacttg aggggcagag tgctgacaga tgaggggcgc acctattgac 6240

atttgagggg ctgtccacag gcagaaaatc agcatttgc aagggtttcc gcccgttttt 6300

cggccaccgc taacctgtct tttaacctgc ttttaaacca atatttataa accttgtttt 6360

taaccagggc tgcgccctgt gcgcgtgacc gcgcacgccg aaggggggtg ccccccccttc 6420

tcgaaccctc ccggcccgct aacgcgggcc tcccatcccc ccaggggctg cgcccctcgg 6480

ccgcgaacgg cctcacccca aaaatggcag cgctggcagt ccttgccatt gccgggatcg 6540

gggcagtaac gggatgggcg atcagcccga gcgcgacgcc cggaagcatt gacgtgccgc 6600

aggtgctggc atcgacattc agcgaccagg tgccgggcag tgagggcggc ggcctgggtg 6660

gcggcctgcc cttcacttcg gccgtcgggg cattcacgga cttcatggcg gggccggcaa 6720

tttttacctt gggcattctt ggcatagtgg tcgcgggtgc cgtgctcgtg ttcgggggtg 6780

cgataaaccc agcgaaccat ttgaggtgat aggtaagatt ataccgaggt atgaaaacga 6840

gaattggacc tttacagaat tactctatga agcgccatat ttaaaaagct accaagacga 6900

agaggatgaa gaggatgagg aggcagattg ccttgaatat attgacaata ctgataagat 6960

aatatatctt ttatatagaa gatatcgccg tatgtaagga tttcaggggg caaggcatag 7020

gcagcgcgct tatcaatata tctatagaat gggcaaagca taaaaacttg catggactaa 7080

tgcttgaaac ccaggacaat aaccttatag cttgtaaatt ctatcataat tgggtaatga 7140

ctccaactta ttgatagtgt tttatgttca gataatgccc gatgactttg tcatgcagct 7200

ccaccgattt tgagaacgac agcgacttcc gtcccagccg tgccaggtgc tgcctcagat 7260

tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct gattacgtgc agctttccct 7320

tcaggcggga ttcatacagc ggccagccat ccgtcatcca tatcaccacg tcaaagggtg 7380

aagcaggctc ataagacgcc ccagcgtcgc catagtgcgt tcaccgaata cgtgcgcaac 7440

aaccgtcttc cggagactgt catacgcgta aaacagccag cgctggcgcg atttagcccc 7500

gacatagccc cactgttcgt ccatttccgc gcagacgatg acgtcactgc ccggctgtat 7560

gcgcgaggtt accgactgcg gcctgagttt tttaagtgac gtaaaatcgt gttgaggcca 7620

acgcccataa tgcgggctgt tgcccggcat ccaacgccat tcatggccat atcaatgatt 7680

ttctggtgcg taccgggttg agaagcggtg taagtgaact gcagttgcca tgttttacgg 7740

cagtgagagc agagatagcg ctgatgtccg gcggtgcttt tgccgttacg caccaccccg 7800

tcagtagctg aacaggaggg acagctgata gaaacagaag ccactggagc acctcaaaaa 7860

caccatcata cactaaatca gtaagttggc agcatcaccc ataattgtgg tttcaaaatc 7920

ggctccgtcg atactatgtt atacgccaac tttgaaaaca actttgaaaa agctgttttc 7980

tggtatttaa ggttttagaa tgcaaggaac agtgaattgg agttcgtctt gttataatta 8040

gcttcttggg gtatctttaa atactgtaga aaagaggaag gaaataataa atggctaaaa 8100

tgagaatatc accggaattg aaaaaactga tcgaaaaata ccgctgcgta aaagatacgg 8160

aaggaatgtc tcctgctaag gtatataagc tggtgggaga aaatgaaaac ctatatttaa 8220

aaatgacgga cagccggtat aaagggacca cctatgatgt ggaacgggaa aaggacatga 8280

tgctatggct ggaaggaaag ctgcctgttc caaaggtcct gcactttgaa cggcatgatg 8340

gctggagcaa tctgctcatg agtgaggccg atggcgtcct ttgctcggaa gagtatgaag 8400

atgaacaaag ccctgaaaag attatcgagc tgtatgcgga gtgcatcagg ctctttcact 8460

ccatcgacat atcggattgt ccctatacga atagcttaga cagccgctta gccgaattgg 8520

attacttact gaataacgat ctggccgatg tggattgcga aaactgggaa gaagacactc 8580

catttaaaga tccgcgcgag ctgtatgatt ttttaaagac ggaaaagccc gaagaggaac 8640

ttgtcttttc ccacggcgac ctgggagaca gcaacatctt tgtgaaagat ggcaaagtaa 8700

gtggctttat tgatcttggg agaagcggca gggcggacaa gtggtatgac attgccttct 8760

gcgtccggtc gatcagggag gatatcgggg aagaacagta tgtcgagcta ttttttgact 8820

tactggggat caagcctgat tgggagaaaa taaaatatta tattttactg gatgaattgt 8880

tttagtacct agatgtggcg caacgatgcc ggcgacaagc aggagcgcac cgacttcttc 8940

cgcatcaagt gttttggctc tcaggccgag gcccacggca agtatttggg caaggggtcg 9000

ctggtattcg tgcagggcaa gattcggaat accaagtacg agaaggacgg ccagacggtc 9060

tacgggaccg acttcattgc cgataaggtg gattatctgg acaccaaggc accaggcggg 9120

tcaaatcagg aataagggca cattgccccg gcgtgagtcg gggcaatccc gcaaggaggg 9180

tgaatgaatc ggacgtttga ccggaaggca tacaggcaag aactgatcga cgcggggttt 9240

tccgccgagg atgccgaaac catcgcaagc cgcaccgtca tgcgtgcgcc ccgcgaaacc 9300

ttccagtccg tcggctcgat ggtccagcaa gctacggcca agatcgagcg cgacagcgtg 9360

caactggctc cccctgccct gcccgcgcca tcggccgccg tggagcgttc gcgtcgtctc 9420

gaacaggagg cggcaggttt ggcgaagtcg atgaccatcg acacgcgagg aactatgacg 9480

accaagaagc gaaaaaccgc cggcgaggac ctggcaaaac aggtcagcga ggccaagcag 9540

gccgcgttgc tgaaacacac gaagcagcag atcaaggaaa tgcagctttc cttgttcgat 9600

attgcgccgt ggccggacac gatgcgagcg atgccaaacg acacggcccg ctctgccctg 9660

ttcaccacgc gcaacaagaa aatcccgcgc gaggcgctgc aaaacaaggt cattttccac 9720

gtcaacaagg acgtgaagat cacctacacc ggcgtcgagc tgcgggccga cgatgacgaa 9780

ctggtgtggc agcaggtgtt ggagtacgcg aagcgcaccc ctatcggcga gccgatcacc 9840

ttcacgttct acgagctttg ccaggacctg ggctggtcga tcaatggccg gtattacacg 9900

aaggccgagg aatgcctgtc gcgcctacag gcgacggcga tgggcttcac gtccgaccgc 9960

gttgggcacc tggaatcggt gtcgctgctg caccgcttcc gcgtcctgga ccgtggcaag 10020

aaaacgtccc gttgccaggt cctgatcgac gaggaaatcg tcgtgctgtt tgctggcgac 10080

cactacacga aattcatatg ggagaagtac cgcaagctgt cgccgacggc ccgacggatg 10140

ttcgactatt tcagctcgca ccgggagccg tacccgctca agctggaaac cttccgcctc 10200

atgtgcggat cggattccac ccgcgtgaag aagtggcgcg agcaggtcgg cgaagcctgc 10260

gaagagttgc gaggcagcgg cctggtggaa cacgcctggg tcaatgatga cctggtgcat 10320

tgcaaacgct agggccttgt ggggtcagtt ccggctgggg gttcagcagc cagcgcttta 10380

ctggcatttc aggaacaagc gggcactgct cgacgcactt gcttcgctca gtatcgctcg 10440

ggacgcacgg cgcgctctac gaactgccga taaacagagg attaaaattg acaattgtga 10500

ttaaggctca gattcgacgg cttggagcgg ccgacgtgca ggatttccgc gagatccgat 10560

tgtcggccct gaagaaagct ccagagatgt tcgggtccgt ttacgagcac gaggagaaaa 10620

agcccatgga ggcgttcgct gaacggttgc gagatgccgt ggcattcggc gcctacatcg 10680

acggcgagat cattgggctg tcggtcttca aacaggagga cggccccaag gacgctcaca 10740

aggcgcatct gtccggcgtt ttcgtggagc ccgaacagcg aggccgaggg gtcgccggta 10800

tgctgctgcg ggcgttgccg gcgggtttat tgctcgtgat gatcgtccga cagattccaa 10860

cgggaatctg gtggatgcgc atcttcatcc tcggcgcact taatatttcg ctattctgga 10920

gcttgttgtt tatttcggtc taccgcctgc cgggcggggt cgcggcgacg gtaggcgctg 10980

tgcagccgct gatggtcgtg ttcatctctg ccgctctgct aggtagcccg atacgattga 11040

tggcggtcct gggggctatt tgcggaactg cgggcgtggc gctgttggtg ttgacaccaa 11100

acgcagcgct agatcctgtc ggcgtcgcag cgggcctggc gggggcggtt tccatggcgt 11160

tcggaaccgt gctgacccgc aagtggcaac ctcccgtgcc tctgctcacc tttaccgcct 11220

ggcaactggc ggccggagga cttctgctcg ttccagtagc tttagtgttt gatccgccaa 11280

tcccgatgcc tacaggaacc aatgttctcg gcctggcgtg gctcggcctg atcggagcgg 11340

gtttaaccta cttcctttgg ttccggggga tctcgcgact cgaacctaca gttgtttcct 11400

tactgggctt tctcagcccc agatctgggg tcgatcagcc ggggatgcat caggccgaca 11460

gtcggaactt cgggtccccg acctgtacca ttcggtgagc aatggatagg ggagttgata 11520

tcgtcaacgt tcacttctaa agaaatagcg ccactcagct tcctcagcgg ctttatccag 11580

cgatttccta ttatgtcggc atagttctca agatcgacag cctgtcacgg ttaagcgaga 11640

aatgaataag aaggctgata attcggatct ctgcgaggga gatgatattt gatcacaggc 11700

agcaacgctc tgtcatcgtt acaatcaaca tgctaccctc cgcgagatca tccgtgtttc 11760

aaacccggca gcttagttgc cgttcttccg aatagcatcg gtaacatgag caaagtctgc 11820

cgccttacaa cggctctccc gctgacgccg tcccggactg atgggctgcc tgtatcgagt 11880

ggtgattttg tgccgagctg ccggtcgggg agctgttggc tggctggtgg caggatatat 11940

tgtggtgtaa acaaattgac gcttagacaa cttaataaca cattgcggac gtttttaatg 12000

tactggggtg gtttttcttt tcaccagtga gacgggcaac agctgattgc ccttcaccgc 12060

ctggccctga gagagttgca gcaagcggtc cacgctggtt tgccccagca ggcgaaaatc 12120

ctgtttgatg gtggttccga aatcggcaaa atcccttata aatcaaaaga atagcccgag 12180

atagggttga gtgttgttcc agtttggaac aagagtccac tattaaagaa cgtggactcc 12240

aacgtcaaag ggcgaaaaac cgtctatcag ggcgatggcc cactacgtga accatcaccc 12300

aaatcaagtt ttttggggtc gaggtgccgt aaagcactaa atcggaaccc taaagggagc 12360

ccccgattta gagcttgacg gggaaagccg gcgaacgtgg cgagaaagga agggaagaaa 12420

gcgaaaggag cgggcgccat tcaggctgcg caactgttgg gaagggcgat cggtgcgggc 12480

ctcttcgcta ttacgccagc tggcgaaagg gggatgtgct gcaaggcgat taagttgggt 12540

aacgccaggg ttttcccagt cacgacgttg taaaacgacg gccagtgaat tgaagcttgg 12600

cgcgccaagc ttggtaccct caggactaga ccatcgtggt taaatgatca agtgcctact 12660

tggcagaatt tctttcgagc agcctcctcc tacaagttgc atttgttgcg cttacgataa 12720

ttgtcaaaga agtaggtaaa ataaagacat gatctctaat attaaggata agattaaaaa 12780

taagtccagg attacccggt cggcccatca attacttgct gacctttgtt gccgtcccac 12840

gacttccatt ttctaaccgt ccattttcat ttgtttttag ctatatttaa tattaatggg 12900

atataaatta taaacattcc tcctcccaaa aaaataagtt taagtaatac tgcaatagac 12960

agtgttttaa gccatgtaat tcagtaaaag ttctttttta ttctgaacct agccctaaaa 13020

aggccatgcg ggtaattagt tcagtcaact gaatatacaa cgttttgaac caaagttaac 13080

atgtacaggc caatagaagt tatttgaccg taagcttagt ctctacattc attcaacgtt 13140

cttgaatcaa agtgacctgt acaggccaat agaagttacc tgaccgtaag cttagtctct 13200

acattcattc ctctgagacg atattctaga agcctgcttt caagtctaaa aggcacaatc 13260

ttttctcctc accacttgtt gaggtactta tgattttaaa gatgaaacat ttttttttact 13320

tttccccttt aatttctttg attttttttt ttcttggtag ttggaagtac ttttcatacc 13380

ctagaaaatc cactgttgat ctttgaaata tcagcaatct ttgaaataat atcagcaacc 13440

acgacaccta ccattctcaa attcactcta taaaagggta aacctttgct tacctctatg 13500

ctcactcaca aggagaacaa acactcatcg gtgctacata accgcggctg caggtcgacg 13560

gatctgtacc cggggatccg tcgatcgttt cgcatgattg aacaagatgg attgcacgca 13620

ggttctccgg ccgcttgggt ggagaggcta ttcggctatg actgggcaca acagacaatc 13680

ggctgctctg atgccgccgt gttccggctg tcagcgcagg ggcgcccggt tctttttgtc 13740

aagaccgacc tgtccggtgc cctgaatgaa ctgcaggacg aggcagcgcg gctatcgtgg 13800

ctggccacga cgggcgttcc ttgcgcagct gtgctcgacg ttgtcactga agcgggaagg 13860

gactggctgc tattgggcga agtgccgggg caggatctcc tgtcatctca ccttgctcct 13920

gccgagaaag tatccatcat ggctgatgca atgcggcggc tgcatacgct tgatccggct 13980

acctgcccat tcgaccacca agcgaaacat cgcatcgagc gagcacgtac tcggatggaa 14040

gccggtcttg tcgatcagga tgatctggac gaagagcatc aggggctcgc gccagccgaa 14100

ctgttcgcca ggctcaaggc gcgcatgccc gacggcgagg atctcgtcgt gacccatggc 14160

gatgcctgct tgccgaatat catggtggaa aatggccgct tttctggatt catcgactgt 14220

ggccggctgg gtgtggcgga ccgctatcag gacatagcgt tggctacccg tgatattgct 14280

gaagagcttg gcggcgaatg ggctgaccgc ttcctcgtgc tttacggtat cgccgctccc 14340

gattcgcagc gcatcgcctt ctatcgcctt cttgacgagt tcttctgagc gggactctgg 14400

ggttcgaaat gaccgaccaa gcgacgccca acctgccatc acgagatttc gattccaccg 14460

ccgccttcta tgaaaggttg ggcttcggaa tcgttttccg ggacgccggc tggatgatcc 14520

tccagcgcgg ggatctcatg ctggagttct tcgcccaccc ccggatcctc tagagtcgac 14580

ctgcaggcgt tcaaacattt ggcaataaag tttcttaaga ttgaatcctg ttgccggtct 14640

tgcgatgatt atcatataat ttctgttgaa ttacgttaag catgtaataa ttaacatgta 14700

atgcatgacg ttatttatga gatgggtttt tatgattaga gtcccgcaat tatacattta 14760

atacgcgata gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc 14820

atctatgtta ctagatcggg aattgccaag ctggcgcgcc ctgcagaccg gtggatccgg 14880

ccggccgatc cggtcatcgg cgggggtcat aacgtgactc ccttaattct ccgctcatga 14940

tcagattgtc gtttcccgcc ttcagtttaa actatcagtg tttgacagga tatattggcg 15000

ggtaaaccta agagaaaaga gcgtttatta gaataatcgg atatttaaaa gggcgtgaaa 15060

aggtttatcc gttcgtccat ttgtatgtgc atgccaacca cagggttccc cagatctggc 15120

gccggccagc gagacgagca agattggccg ccgcccgaaa cgatccgaca gcgcgcccag 15180

cacaggtgcg caggcaaatt gcaccaacgc atacagcgcc agcagaatgc catagtgggc 15240

ggtgacgtcg ttcgagtgaa ccagatcgcg caggaggccc ggcagcaccg gcataatcag 15300

gccgatgccg acagcgtcga gcgcgacagt gctcagaatt acgatcaggg gtatgttggg 15360

tttcacgtct ggcctccgga ccagcctccg ctggtccgat tgaacgcgcg gattctttat 15420

cactgataag ttggtggaca tattatgttt atcagtgata aagtgtcaag catgacaaag 15480

ttgcagccga atacagtgat ccgtgccgcc ctggacctgt tgaacgaggt cggcgtagac 15540

ggtctgacga cacgcaaact ggcggaacgg ttgggggttc agcagccggc gctttactgg 15600

cacttcagga acaagcgggc gctgctcgac gcactggccg aagccatgct ggcggagaat 15660

catacgcatt cggtgccgag agccgacgac gactggcgct catttctgat cgggaatgcc 15720

cgcagcttca ggcaggcgct gctcgcctac cgcgatggcg cgcgcatcca tgccggcacg 15780

cgaccgggcg caccgcagat ggaaacggcc gacgcgcagc ttcgcttcct ctgcgaggcg 15840

ggtttttcgg ccggggacgc cgtcaatgcg ctgatgacaa tcagctactt cactgttggg 15900

gccgtgcttg aggagcaggc cggcgacagc gatgccggcg agcgcggcgg caccgttgaa 15960

caggctccgc tctcgccgct gttgcgggcc gcgatagacg ccttcgacga agccggtccg 16020

gacgcagcgt tcgagcaggg actcgcggtg attgtcgatg gattggcgaa aaggaggctc 16080

gttgtcagga acgttgaagg accgagaaag ggtgacgatt gatcaggacc gctgccggag 16140

cgcaacccac tcactacagc agagccatgt agacaacatc ccctcccccc ttccaccgcg 16200

tcagacgccc gtagcagccc gctacgggct ttttcatgcc ctgccctagc gtccaagcct 16260

cacggccgcg ctcggcctct ctggcggcct tctggcgctc ttccgcttcc tcgctcactg 16320

actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agctcactca aaggcggtaa 16380

tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 16440

aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg ctccgccccc 16500

ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat 16560

aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc 16620

cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt ttccgctgca 16680

taaccctgct tcggggtcat tatagcgatt ttttcggtat atccatcctt tttcgcacga 16740

tatacaggat tttgccaaag ggttcgtgta gactttcctt ggtgtatcca acggcgtcag 16800

ccgggcagga taggtgaagt aggcccaccc gcgagcgggt gttccttctt cactgtccct 16860

tattcgcacc tggcggtgct caacgggaat cctgctctgc gaggctggcc ggctaccgcc 16920

ggcgtaacag atgagggcaa gcggatggct gatgaaacca agccaaccag gaagggcagc 16980

ccacctatca aggtgtactg ccttccagac gaacgaagag cgattgagga aaaggcggcg 17040

gcggccggca tgagcctgtc ggcctacctg ctggccgtcg gccagggcta caaaatcacg 17100

ggcgtcgtgg actatgagca cgtccgcgag ctggcccgca tcaatggcga cctgggccgc 17160

ctgggcggcc tgctgaaact ctggctcacc gacgacccgc gcacggcgcg gttcggtgat 17220

gccacgatcc tcgccctgct ggcgaagatc gaagagaagc aggacgagct tggcaaggtc 17280

atgatgggcg tggtccgccc gagggcagag ccatgacttt tttagccgct aaaacggccg 17340

gggggtgcgc gtgattgcca agcacgtccc catgcgctcc atcaagaaga gcgacttcgc 17400

ggagctggtg aagtacatca ccgacgagca aggcaagacc gagcgccttt ccgacgctca 17460

<210> 7

<211> 24
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide primer

<400> 7
gtcccaatcc accatgcaca tcag 24

<210> 8
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide primer

<400> 8
ccctcgacaa atgcaaaatg tatc 24

<210> 9
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide primer

<400> 9
gatccgctat ggatctttta tc 22

<210> 10
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide primer

<400> 10
atctaatcgt tctttctttg ac 22

<210> 11
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide primer

<400> 11
gcgccgctat ttactcga 18

<210> 12
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide primer

<400> 12
tttctctggc gtgccgcc 18